

VOLUME 21 NUMBER 4 DECEMBER 1958

SOCIOMETRY

A JOURNAL OF RESEARCH
IN SOCIAL PSYCHOLOGY

Articles by

DE VOS and MINER
WATSON
RETTIG, JACOBSON, DESPRES,
and PASAMANICK
COX and KRUMBOLTZ
HOFFMAN
RIECKEN
HAMBLIN

Published by

THE AMERICAN SOCIOLOGICAL SOCIETY

SOCIOMETRY

EDITOR

LEONARD S. COTTRELL, JR.

ASSOCIATE EDITORS

ROBERT F. BALES	NATHANIEL L. GAGE
HERBERT BLUMER	FREDERICK MOSTELLER
JOHN A. CLAUSEN	THEODORE M. NEWCOMB
LEON FESTINGER	HERBERT A. SIMON
NELSON N. FOOTE	FRED L. STRODTBECK

EDITORIAL CONSULTANTS

PETER M. BLAU	DUNCAN MACRAE, JR.
EDGAR F. BORGATTA	JAMES G. MARCH
ORVILLE G. BRIM, JR.	PHILIP J. MCCARTHY
ROGER W. BROWN	THEODORE R. SARBIN
LAUNOR F. CARTER	STANLEY SCHACHTER
SANFORD M. DORNBUSCH	WILLIAM A. SCOTT
FRED E. FIEDLER	SHIRLEY A. STAR
ERVING GOFFMAN	GUY E. SWANSON
ALEX INKELES	JOHN W. M. WHITING
WILLIAM W. LAMBERT	ROBIN M. WILLIAMS, JR.
RICHARD T. LAPIERE	ROBERT F. WINCH

EXECUTIVE OFFICER

MATILDA WHITE RILEY

Published four times a year

March, June, September, December

by

THE AMERICAN SOCIOLOGICAL SOCIETY

at The Boyd Printing Co., Inc., 372-378 Broadway, Albany 7, N. Y.

Sociometry was founded in 1937 by J. L. Moreno

Communications for the Editors and all manuscripts should be addressed to Leonard S. Cottrell, Jr., Russell Sage Foundation, 505 Park Avenue, New York 22, New York. Business communications should be addressed to Matilda White Riley, The American Sociological Society, New York University, Washington Square, New York 3, New York. Copyright © 1958 by THE AMERICAN SOCIOLOGICAL SOCIETY
Entered at the Post Office at Albany, N. Y., as second-class matter.

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

VOLUME 21 . NUMBER 4 . DECEMBER 1958

CONTENTS

- Algerian Culture and Personality in Change
George De Vos and Horace Miner 255
- A Formal Analysis of Sociable Interaction.....*Jeanne Watson* 269
- Rating Response Set as a Function of Objective Status Criteria
*Salomon Rettig, Frank N. Jacobson, Leo Despres, and
Benjamin Pasamanick* 281
- Racial Bias in Peer Ratings of Basic Airmen
John A. Cox and John D. Krumboltz 292
- Similarity of Personality: A Basis for Interpersonal Attraction?
L. Richard Hoffman 300
- The Effect of Talkativeness on Ability to Influence Group Solutions
of Problems.....*Henry W. Riecken* 309
- Leadership and Crises.....*Robert L. Hamblin* 322
- Annual Author Index—Volume 21..... iii

SOCIOMETRY

EDITORIAL POLICY

Sociometry is concerned with the entire range of interests and problems represented by research in social psychology. It is the policy of the editors to seek those manuscripts for publication which represent the significant research interests of investigators who are concerned with giving the field of social psychology theoretical structure and reporting research which is clearly focused, well designed, and competently conducted.

While social psychology is presently regarded by most as a field with indeterminate boundaries, it has as its central focus the investigation of the processes and products of social interaction at the interpersonal, intrapersonal, intergroup, and intragroup levels and the development of significant generalizations therefrom. In keeping with the more general meaning of the name of the journal emphasis will be placed on measurement of social behavior. However, this emphasis does not exclude the acceptability of good articles which must rely upon qualitative materials and analyses.

The editors and editorial consultants can be expected to subject manuscripts to rigorous criticism and screening according to the best standards of scientific research and at the same time avoid a sterile orthodoxy which would stultify the communication of creative ideas at the growing edge of the science. Thus the journal will strive to be flexible in its response to the publication needs of its contributors.

It is the intention of the editors to avoid any tendency toward professional provincialism and to invite contributions from any sector of the scientific community which promise to further the objectives of the journal.

Annual *subscription* rate \$9.00; special rate to members of THE AMERICAN SOCIOLOGICAL SOCIETY \$4.50. Single issues \$2.50.

Specifications for Manuscript Preparation for Sociometry may be obtained from the Editorial Office.

Back Issues dated 1937 through 1955 should be ordered from the former publisher, Beacon House, Beacon, New York.

Notice of *change of subscriber's address* should be sent to THE SOCIETY four weeks in advance, indicating the old as well as the new address.

Algerian Culture and Personality in Change¹

GEORGE DE VOS, *University of California*

HORACE MINER, *University of Michigan*

The Algerians we shall consider are known to the French as "Arabs" and, in fact, a thousand years of residence in North Africa has not obscured their racial and cultural ties to Arabia. Nevertheless, since the French occupied Algeria over a century ago, the Arabs have been subjected to marked acculturative influences, particularly in the French-dominated cities. The research reported here is specifically addressed to understanding the nature of the relation between Arab culture and personality under the impact of French urban influences.

In considering Arab culture, we have departed from the traditional anthropological method of describing culture traits in terms of the usual, average, or preferred behavior. Such descriptions possess a useful economy for some purposes, but they tend to give a false impression of homogeneity. Particularly in an investigation of the relation between psychological and cultural phenomena, it seems desirable to express the cultural forms in terms of distributions of individually varying behavior. When this is combined with individual personality data, each person, as well as the culture group, becomes a meaningful item of analysis.

For our purposes, "personality traits" will refer to those basic psychological structures which are expressed in the generalized tendency of an individual to relate himself to various life situations in some characteristic fashion. Operationally defined, these traits will be those which emerge from the analysis of Rorschach protocols in conjunction with personal observations, but not depth interviews.

The basic design of the study involves the synchronic comparison of Arabs living in a small oasis in the northern Sahara with other Arabs who were born and reared in this oasis but who left it some time after adolescence to take up residence in the city of Algiers. These characteristics of the two groups minimize the degree of cultural difference which might be expected between them but maximize the probability of homogeneous origins of the two groups. The importance of the latter is obvious in a study in which

¹ De Vos is responsible for the scoring of the Algerian Rorschach protocols and for their interpretation. Miner collected the cultural data and administered the tests during field work made possible by a grant under the Fulbright Act and another from the Horace H. Rackham Fund. The authors wish also to acknowledge the assistance of Bernard Berk, Akira Hoshino, Takao Sofue, and Mayumi Taniguchi in the processing and statistical analysis of the primary data.

group differences are to be interpreted as due to acculturative change. The possibility of selective urban migration cannot be ignored, but comparison of the two groups on the basis of a substantial body of family background data provides assurance as to their similarity of origin. The only selective factor apparently operating is that of economic pressure, but the necessity for migration out of the oasis is so great and so general that even the families that are economically more secure contribute sons to the urban movement. Such differences in background as were found were controlled in the comparative analysis.²

The Setting

The date-growing oasis of Sidi Khaled has a resident population of 5,300. The population has doubled in the last century, but further expansion of the agricultural economy is stifled by lack of water. As a result, two thirds of the men between the ages of twenty and fifty must now seek unskilled work in the cities of France and North Africa. A few of those who go to the city return after failing to find work or after accumulating a little capital. These returnees exert some acculturative influence in the oasis, and a French school has sporadically exposed a small group of boys to the three R's. Despite these influences, life for most of the oasis dwellers continues in terms of traditional Arab culture.

Those who move to Algiers find themselves in a new physical world, peopled with what appears to the newcomers as domineering French and with Arabs who have already altered many of their beliefs as a result of long interaction with the French. The secluded courtyard of the oasis gives way to the crowded apartment. Opportunities to work are controlled by the French or by relatively acculturated Arabs, and there are both economic and status advantages to adopting Franco-Arab city ways. To find a role in the life of the city, the migrant is inescapably forced, or lured, into adopting its patterns of life.

Within this oasis-urban comparative framework, we shall consider the cultural and Rorschach data³ collected in 1950 from 64 Arab men, largely between the ages of twenty and fifty. Twenty of these men, to whom we shall refer as the "oasis group," either have never been outside of the oasis setting or have had less than four months' residence in the city. The "urban

² A full report of the investigation is in process of publication (9).

³ The Rorschach protocols were transcribed from tape recordings obtained by Miner in Algeria. The protocols were scored independently by two research assistants according to a standardized system (4). These scorings were then checked by De Vos and any discrepancies in the previous two scorings were resolved. The tape recordings allowed for a full reference to the detailed inquiry thus assuming a relatively high level of unanimity in scorer reliability.

group" consists of 28 men who have lived more than a quarter of their lives in the city. All have resided there longer than five years and most for more than twenty. The 16 Arabs who had amounts of contact intermediate between these oasis and urban groups are excluded from the comparisons to be considered, but they figure in the over-all analysis of Arab personality.

An attempt to work with subjects selected on a random basis proved utterly impractical. The oasis and urban groups, therefore, cannot be said to be representative. Both groups of subjects were, however, chosen so as to provide a wide range of age, occupation, and prestige. Despite the informal method of selection, the oasis and urban samples are similar in terms of their basic demographic characteristics. Each of the subjects was interviewed concerning the same areas of culture and a Rorschach test was administered to each.

Some Psychological Variables Related to Cultural Change

For purposes of this presentation ⁴ we shall deal mostly with those Rorschach variables that showed significant relations to specific cultural beliefs and opinions. The measures that showed the closest relationship to cultural traits were not the Rorschach variables usually scored in quantitative analysis but (a) over-all indices of rigidity and maladjustment, using a method of assessment developed by Seymour Fisher (7), and (b) a system for scoring the affective implications of the content of Rorschach responses developed by De Vos.⁵ These scoring systems have been used previously with both American and non-Western data (4, 5, 6). There were certain modal psychological patterns found in the protocols that did not differentiate between the urban and rural groups but were highly characteristic for the group as a whole.

There were, for example, characteristic thought patterns in evidence that conformed remarkably to a pioneering Rorschach study by Bleüler and Bleuler done with Moroccans twenty years ago (3). Many individuals tended to create percepts full of arbitrary juxtapositions on the one hand or arbitrary discriminations on the other. The records were replete with modes of thought suggestive of what would be diagnosed in Western records as obsessive-compulsive character defenses. In the more extreme cases the type of illogical juxtapositions brought forward in the organization of the responses, as well as the modes of rationalizing the responses given, were strikingly similar to

⁴ A more detailed presentation of the statistical treatment of all the Rorschach data was prepared for the full report now in process of publication (9).

⁵ In this system the affective implications of the content are scored according to a number of discrete categories of hostility, anxiety, etc. These categories are combined for quantitative treatment into over-all indices of hostility, anxiety, body preoccupation, dependency, positive, miscellaneous, and neutral content. See De Vos (4) for the criteria used in scoring according to this system.

those found in paranoid records in Western groups. An examination of certain individuals with sufficient background evidence available demonstrated that the patterns of thought indicated in the Rorschach were indeed related to behavioral evidence. In a number of instances the test seemed to differentiate as well between Arabs as it does between Americans, assuming sufficient knowledge of the cultural background. In the main, the determinants of responses and the location used were similar for both the rural and urban groups. However, certain differences in the use of color, shading, and inanimate movement distinguished the urban-dwelling Casbah group from those of the oasis. The picture is a consistent one in which the urban group inferentially demon-

TABLE 1

Rorschach Indices of Rigidity and Maladjustment in Arab and American Samples

	N	Arab			American		
		Oasis	Urban	Total sample	Normal	Neurotic	Schizophrenic
		20	28	64	60	30	30
Rigidity	Mean	51.0*	46.3*	50.1*	27.7	30.8	32.2§
	S.D.	10.3†	20.2†	14.0	15.3	15.3	16.8
Maladjustment	Mean	63.3*	61.8*	62.2*	34.0	65.8*‡	80.6*‡
	S.D.	19.8	26.1	20.2	16.2	31.7	23.8

Key to levels of significance:

* Difference from American Normal significant (at .05 level).

† Oasis-urban Arab difference in standard deviations on the Rigidity Scale significant.

‡ Neurotic-schizophrenic American difference significant.

§ Schizophrenic-normal American difference significant.

strated more tendencies toward diffuse anxiety, less conscious control over affective reactions, and more tension over the control of impulses.

On over-all assessment of rigidity, using the Rigidity scale, the Arabs scored extremely high. The group as a whole obtained a mean score over one standard deviation above the mean of a sample of American normals. This mean was considerably beyond measures of rigidity as applied to maladjusted American groups (see Table 1).⁶ On the maladjustment scale that was applied to the records, the mean score obtained for the group as a whole was directly

⁶ The American samples cited in Table 1 for comparative purposes are derived from the normative files of Dr. S. J. Beck and have been used extensively by him and others in previous research (2). De Vos employed them previously in a comparison with acculturating Japanese Americans (5, 6). Their characteristics are described in some detail elsewhere (5). They correspond sufficiently in age and other considerations to be used meaningfully as a control group in citing the Algerian data.

comparable to that of an American neurotic sample and over two standard deviations above the mean of the American normal group.

In a forthcoming book by Miner and De Vos (9) there is a discussion of the implications of the high rigidity and maladjustment scores to an understanding of Algerian Arab personality. Briefly summarizing, these scores measure imbalances in the use of certain personality defenses or control mechanisms. Their psychodynamic implications are the same for Arabs or Americans, but cultural modalities may allow for the presence of certain types of rigidity or imbalance in psychological defenses in one setting that would lead to more behaviorally manifest signs of maladjustment in another setting.

When the rigidities and psychic defenses of the Arabs are in line with modal cultural practices the rigidity and maladjustment suggested by the Rorschach is behaviorally "latent" and nondisruptive, and behavior is perceived as acceptable to other Arabs. If more mature and flexible adaptive mode should be called forth with change such as in an acculturative situation, the inflexibility and maladaptive nature of the defensive structure can in many instances become more manifest to the observer.

While it does not follow that the Algerians averaging such high scores in these measures be considered neurotic as a group, behavioral evidence was sufficient in a number of cases to indicate intrapsychic disturbances in certain capacities to adjust that was in line with the Rorschach evidence. In other cases, however, no overt maladjustment was perceivable to the anthropologist nor noted by others.

The maladjustment and rigidity measures used are not completely independent as to the variables scored. There is some tendency for individuals with excessively high rigidity of necessity to show up high on the maladjustment scale as well. However, this sort of relationship between the two scales is not always in evidence. In the Arabs these scales tended to co-vary only in the case of the urban group. It was in some of the Casbah records that the maladjustment scores were related especially to the presence of excessively inflexible ego defenses, suggesting that the individuals in question were not readily adaptive to changes in the environment.

One might raise the question, at least in regard to these over-all results with Algerians on a maladjustment score, of whether the type of personality integrations found modal for Arabs are sufficiently different from those of Americans or French to involve special problems in achieving a ready adjustment to Western life.

No significant difference was found between the oasis and urban groups in either mean rigidity or mean maladjustment. In the city, however, the mean rigidity score is somewhat lower and shows significantly greater variability due to the appearance in a minority of individuals (including those be-

haviorally more successful in acculturation) of lower rigidity scores. These records are more comparable to American norms. The number of these records is not sufficient, however, to lower the mean significantly below that of the oasis group, since the city group also has a number of individuals with high rigidity that co-varies with high maladjustment. The oasis group's rigidity and maladjustment scores did not co-vary. The most rigid records were in no case among the highest in maladjustment.

Turning to the analysis of symbolic content of the Rorschach protocols, the proportion of unpleasant content was found to be significantly higher among the city Arabs, particularly in the content categories indicative of body preoccupation or hostility. These results, as was indirectly indicated by the results with color and shading and inanimate movement determinants, suggest that urban residence produces increased psychic stress for the Arabs (see Tables 2a and 2b). Two patterns of adjustment which appear with some frequency in the city are only rarely found in the oasis. One centers around greater rigidity and internalization of aggression suggested by anatomical preoccupations. Such records contributed heavily to the high mean maladjustment score found in the urban group. Another pattern suggests the development of a more complex, flexible ego, but one which must cope with heightened internalized tendencies for aggressive hostility to be overtly expressed. In this latter pattern a considerable amount of tension over hostility is projected outward and the environment tends to be peopled with hostile and dangerous human forces. It is reasonable to believe from the Rorschach evidence that the degree of threat which the French represented to these Algerians was a reflection of intrapsychic stress as well as the continuing effect of objective political and economic issues.⁷

Personality Variables Related to Specific Cultural Beliefs

These characteristics of the urban Arab as compared to those living on an oasis are in some cases related to changes in culture in a direct manner. The ensuing discussion will, therefore, consider the interrelationships among three types of variables: (a) the personality characteristics just discussed; (b) urban contacts; and (c) cultural beliefs associated with the seclusion

⁷ Goldfarb (8) reports in a study of the American Negro the presence of anatomical content and responses in which the human body is perceived mutilated in some form in 88 per cent of the cases included in Kardiner and Ovesey's study, *The Mark of Oppression*. Similar responses were notably present in Japanese immigrants to America (5) in contrast to those living in Japan. Abel and Hsu, although using a different interpretation, report anatomy responses prevalent in male Chinese-Americans (1). The similarity of this sort of symbolic content in such widely disparate minority groups attests to certain possible common intrapsychic coping mechanisms related to the environmental stress of a depreciated minority status.

TABLE 2a
*Comparison of Indices of Affective Symbolism **

N	Arab				American			
	Oasis		Urban		Normal		Neurotic	
	20		28		60		50	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Hostility %	6.8	8.6	10.0	7.9	9.4	7.1	11.6	6.9
Anxiety %	19.1	14.2	18.7	14.9	13.8	7.6	22.6*	11.8
Bodily preoccupation %	2.0*	3.6	10.1*	19.0	4.1	4.9	9.3*	12.8
Total unpleasant	27.9*	17.5	38.8*	23.8	27.3	12.0	43.5*	13.2
Dependent %	7.1	7.1	6.1	8.0	4.6	4.6	7.0	7.4
Positive %	13.2	10.7	9.4	9.2	17.2	10.9	11.9*	9.8
Neutral %	49.3	19.6	43.0	37.5	49.3	15.6	45.3 ^b	19.2
							34.6*	19.1

* Comparison with American Normal significant (at .05 level).

* Comparison between oasis and urban Arabs significant.

^b Comparison between American neurotics and schizophrenics significant.

Categories do not total 100% because of omission of "Miscellaneous."

TABLE 2b

Presence of Significantly High Scores in Hostility or Body Preoccupation

	Hostility	Body preoccupation	Either or both
Oasis (N=20)	3	1	3
Casbah (N=28)	7	6	12
Total Arab (N=64)	10	7	15

P=.05

*Presence of Mutilated or Distorted Humans, or Tension Responses
(Such as Explosions)*

Scored under hostile symbolism	Humans mutilated or distorted (Hsm and Hhad) **	Tension response (Hhat) **	Either or both
Oasis (N=20)	6	1	7
Casbah (N=28)	13	8	17
Total Arab (N=64)	28	10	34

P=.05

* 1 S.D. above mean of normal Americans.

Hostile 16%+.

Body preoccupation 10%+.

** See "A Quantitative Approach to Affective Symbolism in Rorschach Responses" (4) for a full description of criteria for scoring content.

of women, the punishment of children, and those connected with certain supernatural beliefs.⁸ See Table 3.

Seclusion of Women

Concerning the seclusion of women, we note that from the time of puberty they are carefully isolated from the sight of men other than their husbands and close kinsmen. A woman's life tends to be narrowly restricted to her home, and when she does leave its confines, she must be veiled and accompanied.

The relation of three seclusion customs to the personality variables was explored, as will be seen from the accompanying table. The degree of veiling required and the terminal age for the seclusion of women both showed a statistically verifiable consistent relationship to rigidity and maladjustment. The more restrictive customs were preferred by Arabs who were very high in these qualities, and, interestingly enough, somewhat low in symbolic indications of hostility and body preoccupation in the content of Rorschach responses. Concerning the acceptability of wife's mother as her chaperone,

⁸ Tests of significance of relationships are based on chi square, using a correction for continuity, or, particularly when theoretical cell frequency was below 5, Fisher's Exact Test was employed and the resultant probability doubled to provide the closest possible approximation of a two-tailed test.

TABLE 3

*Levels of Significance of Relationships between Rorschach and Cultural Variables **

	High rigidity	High maladjustment	Medium or high in positive content	High in hostility and body preoccupation
Cleanliness				
Meticulous or clean	(-.10)	(-.10)	—	—
Seclusion				
Requires veiling of one or both eyes	(.10)	.05	—	(-.10)
Forbids wife to go out with her mother	—	—	—	—
Wife's seclusion ends at 56 years or older	.001	.05	—	(-.15)
Punishment				
Beats severely or prefers beating	.05	—	—	(-.15)
Uses isolation	—	—	—	—
Uses food deprivation	-.05	(-.10)	.05	—
Supernatural				
Uses charms	—	—	.05	-.05
Believes in genii	—	—	—	(-.15)
Believes in power of <i>sakkara</i>	—	—	—	-.05
Protects against or has been affected by evil eye	—	—	.05	-.05
Interprets his dreams	—	—	—	—
Believes in power of <i>guezana</i>	—	—	—	—

* Total sample analyzed with two-tailed tests of significance. Minus signs indicate an inverse relationship between the variables.

however, there is a striking lack of relationship to the psychological variables, despite the seeming similarity of the personality implications of the three seclusion customs.

If we place the evidence in the general context of what was learned about Arab personality, we conclude that the relatively rigid maladjusted Arab, whose psychic problems are inferentially markedly sexual in origin, imposes strict controls on his women. In the tendency for an inverse relationship to exist between severity of restriction and hostility or body preoccupation symbolism, we see evidence that the rigid, maladjusted Arab who follows such restrictive practices does not tend to develop more internalized and intrapsychic con-

flicts over the handling of his aggressions. In a sense, they are "taken out on the women" in acceptable cultural ways which produce no overt or even covert hostile tensions.

The above cultural expression of personality characteristics is based on analysis of the total Arab sample. Beyond demonstrating its existence, it should be revealing to see how these apparently functional links between culture and personality operate under acculturative influences by contrasting urban and rural groups.

Considering veiling norms first, we note that the shift to the more liberal requirement of allowing the exposure of both eyes is so general in the city that only four of the urban Arabs continued to insist on more complete covering. The demonstrated personality implications of differences in veiling practice are, therefore, essentially an expression of the oasis situation.

The relationship of maladjustment to veiling is significant in the oasis group taken alone where there is still possibility of measuring differences in attitudes. It is revealing, however, to examine the urban "traditionalists," despite their small number. Of these four more restrictive Arabs, all are high in rigidity, three are high in maladjustment, and three are low in both hostility and body preoccupation, paralleling the oasis findings.

It is meaningful to ask if the change in cultural norms occurred despite the apparent psychic "utility" of the oasis norms in personality integration, or whether the changes in cultural attitudes are paralleled by personality shifts attendant upon urban residence. We have already seen that over-all indices of rigidity and maladjustment do not change in the city, but Rorschach content symbolizing hostility and body preoccupation increases. The total evidence indicates, therefore, that change in attitudes toward veiling goes on despite its significant relationship in the oasis to certain personality traits and basic attitudes toward women which do not change with movement to the city. The urban increase in hostility symbolism or anatomical responses is consistent with the decline in severity in attitudes toward veiling. As will be further discussed below in relation to supernatural belief, it would not be unreasonable to assume that the external social pressures of the Casbah affect cultural beliefs. With urbanization, certain changes in psychological integration must occur. The acculturating individual finds himself bereft of culturally condoned beliefs and practices that would help circumvent intra-psychic conflict. The person who is more stripped of reassuring religious and social beliefs is also more exposed to the French world and what is perceived as its threatening domination.

There is less change in prevalence of attitudes between oasis and Casbah as to the age at which a woman may come out of seclusion. The relationship

with content symbolism rigidity and maladjustment scores,⁹ nevertheless, is similar to the results with veiling, namely those who are most rigid and maladjusted are more severe in adherence to custom, whereas those who give more hostile or body preoccupation content are prone to be less severe.

Punishment of Children

The psychological implications of the methods employed in punishing children are markedly varied. The use of isolation as a punishment shows no relation to the personality measures employed. The use of severe physical punishment is characteristic of the most rigid Arabs and of those who do not tend to show marked hostility or body preoccupation. But unlike seclusion customs, such beating of children shows no relation to maladjustment. Beating does not become less severe in the city, nor does rigidity decrease, but the expression of preference for such beating as the best form of punishment does decline. The Arab seems to beat his sons through rigid adherence to the pattern learned on the receiving end in childhood. The father's domination of his children is not altered by contact with the French, but he is probably not so likely to express, at least to a non-Arab, what he knows will be taken as an unduly punitive position.

The most rigid Arabs, however, are significantly less likely to punish children by depriving them of a meal, limiting their food, or even delaying their meals. The explanation here seems to be that beating is the usual form of punishment and only the less rigid Arabs are likely to depart from the pattern and employ other means. Interestingly, it is also those who tend to be better adjusted and who give more positively toned affective material in the Rorschach who punish by food deprivation. Such Arabs show no less tendency to use physical punishment. It is as though those who use deprivation to punish their children do so with the feeling that "it is for their own good." Despite the specific association of less rigid personality factors with food deprivation, use of this kind of punishment increases generally in the city.

Supernatural Beliefs

Rigidity and maladjustment, which showed such striking relationships to the preceding culture traits, are not related to any of the six supernatural traits analyzed. However, the relation of maladjustment to belief or disbelief in fortune-telling *guezzanas* was significantly different in the city and oasis. In the desert eleven of the twelve disbelievers are very high in maladjustment.

⁹ In the oasis "age out of seclusion" is significantly related to rigidity at the .05 level and in the city it almost reaches this level. In the oasis the relationship to maladjustment does not reach significance, but in the city it is significant.

It is possible that a religious interdiction of fortune-telling is responsible for the tensions associated with this custom, but the direction of the relationship would have been hard to predict.

The most marked personality concomitant of the supernatural beliefs rated is the inverse relationship which exists between the intensity of four of the beliefs and indications of hostility or body preoccupation. As with seclusion customs, approved belief in supernatural forces seems to function to obviate the formation of aggressive sets either turned outward toward others or directed inward against the self that appear in Rorschach content symbolism. Customs concerning charms, the evil eye, and love magic of the *sahhara* show this pattern at statistically significant levels,¹⁰ while belief in *genii* shows a trend of similar nature. The only beliefs which diverge from this pattern are the two involving prognostication. As a concomitant of the inverse relationship of four of the traits to types of unpleasant Rorschach content, two supernatural beliefs show significant direct relations to positive content in the test.

One may provisionally conclude on the basis of these results that the personality implications of the supernatural traits hold without regard for acculturative changes. Customs concerning charms, *genii*, and *sahharas* wane significantly in the city, where there is a concomitant increase in evidence of hostility and bodily preoccupation symbolism. But concern over the evil eye, which has similar personality implications, is not altered by urban contact. In fact, all the evidence we have presented indicates that culture change goes on in response to social pressures, without much regard for the previous personality implications of the traits involved. It would certainly have been impossible to predict which culture traits would change on the basis of our knowledge of their psychological significance for the individual in the oasis and our knowledge of the personality shifts attendant upon urban contact. There is probably some sort of threshold beyond which given personality types cannot adjust to certain cultural behaviors. Our investigation, however, seems to indicate a marked tendency for personality predispositions and cultural configurations to develop new kinds of equilibria during acculturation.

It is important to point out also that half of the urban group came to the city when they were fourteen years of age or older and three quarters were over nine years old. The changes which they underwent, both culturally and psychologically, demonstrate that the effect of early training and experience

¹⁰ We have seen that fortune-telling is significantly and differently related to maladjustment in the oasis and city. Similarly, the use of dream interpretation is directly related to low anxiety in the city and the city differs from the oasis in this regard. The only other significant finding with respect to supernatural traits in the subsamples was the inverse relation of use of charms to hostility and body preoccupation in the urban group taken above.

is not conclusive in personality formation. It is the continuity of influences through life, and not just the impact of early influences alone, which makes men as they are.

In summary, then, the evidence from Rorschach protocols, when viewed in relation to seclusion practices, discipline of children, and religious beliefs, shows a consistent pattern, namely that attenuation of traditional beliefs in the urbanized Arabs is related to increasing intrapsychic tensions that are expressed in symbolic form in Rorschach content in a number of individuals. The minority position of the more acculturated urbanized Arab is reflected in internalized personality adaptations in which the social environment is more directly experienced as hostile and threatening. He must cope with the implications of this psychological set intrapsychically. Direct expression of reactive hostility was not readily possible. Attention is more directly focused on feelings of being directly oppressed in relation to the dominant French. Those who adhere more tenaciously to traditional beliefs may demonstrate greater rigidity in certain instances and score higher on maladjustment indices, but in adhering to social and religious beliefs they are not forced into patterns of adaptation that cause them to experience social relationships as directly involving a great degree of personal threat coming from their own projected hostility. The fact that analyses of materials on urban American Negroes living in New York (8) show similar Rorschach patterns of personality integration is highly suggestive for understanding the effects of minority group status on individual integration where pronounced rejection by the dominant group makes actual assimilation impossible.

Manuscript received: February 17, 1958

Revised manuscript received: July 29, 1958

George De Vos
School of Social Welfare
University of California
Berkeley 4, California

REFERENCES

1. Abel, T. M., and F. L. K. Hsu, "Some Aspects of Personality of Chinese as Revealed in the Rorschach Test," *Rorschach Research Exchange and Journal of Projective Techniques*, 1949, 13, 285-301.
2. Beck, S. J., et al., "The Normal Personality as Projected in the Rorschach Test." Published as a separate and in the *Journal of Psychology*, 1950, 30, 241-298.
3. Bleuler, M., and R. Bleuler, "Rorschach's Ink Blot Test and Racial Psychology: Peculiarities of Moroccans," *Character and Personality*, 1935, 4, 97-114.
4. De Vos, G., "A Quantitative Approach to Affective Symbolism in Rorschach Responses," *Journal of Projective Techniques*, 1952, 16, 133-150.

5. De Vos, G., "A Comparison of the Personality Differences in Two Generations of Japanese by Means of the Rorschach Test," *Nagoya Journal of Medical Science*, 1954, 17.
6. De Vos, G., "A Quantitative Rorschach Assessment of Maladjustment and Rigidity in Acculturating Japanese Americans," *Genetic Psychological Monographs*, 1955, 52, 51-87.
7. Fisher, S., "Patterns of Personality and Some of Their Determinants," *Psychological Monographs*, 1950, 64.
8. Goldfarb, W., "The Rorschach Experiment," in A. Kardiner and L. Ovesey, *The Mark of Oppression*, New York: Norton, 1951.
9. Miner, H., and G. De Vos, *Oasis and Casbah: Algerian Culture and Personality in Change*. In preparation for publication.

A Formal Analysis of Sociable Interaction¹

JEANNE WATSON, *University of Chicago*

Social psychology has made great strides in the study of interaction as it occurs in a face-to-face group. There has been relatively little effort, however, to develop a theoretical framework that would distinguish work-oriented interaction from other types of interaction and specify the characteristic forms and processes associated with each type.

FORMAL CHARACTERISTICS OF THREE STYLES OF INTERACTION

Our recent studies of sociable interaction have resulted in the formulation of three different models for the description of face-to-face interaction. Each model is to be regarded as an ideal type which seldom occurs in its pure form. The three types of interaction have been called, roughly, work-oriented, familial, and sociable. In each case, formulation of the characteristics of the particular type of interaction began with an analysis of the functions of the social institution which gave the type its name. Thus, for example, we tried to think about what kind of interaction would make it possible for the family to perform its unique functions and called this kind of interaction familial. Similarly, the kind of interaction which would make work possible was called work-oriented, and the kind of interaction which would facilitate sociability was called sociable.

Integration of Individuals with Society

We may begin by outlining the differences among the three types of interaction as given by five different criteria. The first criterion has to do with the way in which the interaction relates individuals to the larger society in which they live. We can say that work involves the production of goods and services which will be useful to society, and, correspondingly, that work interaction involves a division of labor, providing for functional integration of individuals with each other and for relating these individuals to society on the basis of productivity.

There are several different statements one could make about how the family

¹ This paper grows out of a research project made possible by research grant 891 from the National Institute of Mental Health, Public Health Service. It was read at the 34th Annual Institute of the Society for Social Research of the University of Chicago, 1957. The ideas presented here have been influenced by staff discussions with David Riesman, Robert J. Potter, Nelson N. Foote, and Reuel Denney; and also by the work of A. T. M. Wilson and Elliott Jaques.

TABLE 1
Three Styles of Interaction

Criterion	Work-oriented interaction	Familial interaction	Sociable interaction
Integration of individuals with society	Functional integration; productivity	Privacy and retreat	Normative integration
Individual need fulfillment	Achievement, competence	Unconditional love and acceptance	Unique personal qualities
Investment of self	Segmental; successes	Problematic or routine	Participate in group-culture-building
Conversational style	Progressive	Reiterative	Dramatization of new experience
Conversational resources	The job to be done	Daily life	Personal interests

relates the individual to society, such as that it provides for reproduction, or for socialization, or for an extended set of kinship ties. For purposes of this analysis, we have chosen to emphasize the fact that the family can provide an opportunity for privacy and for retreat from society. The familial style of interaction, then, is one which proceeds from the assumption that the outside world is in some sense alien, with the family, or pseudo-family, serving as sanctuary for the individual.

In considering sociability we have also been selective in emphasizing only one part of what actually goes on, namely, the development of shared values and shared definitions of reality. The sociable style of interaction is viewed as one which provides for normative integration of the individual with society.

The Satisfaction of Personal Needs

A second criterion used in setting up specifications for the three types of interaction concerns the contribution of each type to the satisfaction of personal needs. Here we assume that work gives an opportunity for achievement; the family, at least theoretically, offers unconditional love and acceptance to each member; and sociability provides an opportunity for each person to experience himself as a unique individual. The same point could be stated differently by saying that individual distinctions at work are essentially hierarchical, based on demonstrated competence. Distinctions within the family are made on the basis of role, such as mother, father, son, brother, daughter,

or sister, but love is given equally to all. It is only in sociability that the emphasis is placed on nonhierarchical qualitative distinctions, with the important thing being an individual's personality and not his achievements or role. In terms of interaction, we can say that the work style emphasizes individual competence and achievement, the familial style emphasizes unconditional acceptance of all persons present, and the sociable style emphasizes the unique personal qualities of each individual.

Investment of Self

A third criterion has to do with the aspect of self which is invested in the interaction. The work style requires segmental presentation of self in which each individual demonstrates his competence and success in a restricted area defined by the job. The familial style tends to emphasize presentation of problematic or routine aspects of the self with the listener's personal interest in the speaker being taken for granted. In sociability, we would suggest that the characteristic style is what may be called culture building, in which individuals join together to create and maintain a world of special meanings—whether these meanings be literary, political, personal, social, or even objectively silly and ridiculous. In this case, the individual dramatizes that part of himself which overlaps with the group culture.

Conversational Style

The fourth criterion which we have used is a stylistic one, having to do with the manner in which topics are discussed. The work style emphasizes making progress, getting things done, completing one task and moving on to the next. Most of the analysis of group problem-solving illustrates this aspect of work interaction. The familial style is essentially repetitive, dealing again and again with the same routine concerns. The sociable style is built around the dramatization of new experience, with the emphasis placed both on novelty and on the entertainment value of events and experiences.

Conversational Resources

The last criterion refers to the subject matter of conversation, or what we have called the conversational resources. Work-oriented interaction focuses on the job to be done; familial interaction focuses on routine events and problematic concerns from the daily lives of the participants; and sociable interaction is more concerned with the special interests and nonroutine concerns of the individuals.

These five distinctions should give a fairly clear picture of what is meant by each of the three styles of interaction. From this point on, we shall con-

sider each type of interaction as an independent entity, divorced from the setting which provided the original defining characteristics. For instance, we want to be able to say that the sociable style of interaction can be found both at work and in the family; and similarly, to say that interaction in sociable settings may follow work or familial styles, as well as sociable ones. Perhaps it would be better if we were to call these interaction styles A, B, and C, and drop the labels of familial, work, and sociable. However, we feel there is still some intuitive meaning to these labels, and we shall continue to use them. We only ask that the reader not take them too literally. When we talk about sociable interaction, we are referring to one particular style of interaction which can occur under a variety of conditions. We are not saying that this style of interaction is restricted to sociable settings, or even that it is characteristic of them.

FACTORS WHICH INFLUENCE INTERACTION IN A SOCIABLE SETTING

It is possible to identify three different kinds of pressure which help to give form to interaction in a sociable setting—hereafter called sociability. Each of these forces can be examined to see whether it tends to emphasize work, familial, or sociable styles. Briefly, these three types of force have to do with historical and social class determinants; with the use of sociability as a defense against individual anxiety; and with the characteristics of sociability as a structural form.

Historical and Class Influences upon Sociability.

Let us consider the historical and class determinants first. The major point here has to do with *democratization*. There was a time when sociability as a social form was pretty well restricted to a leisured aristocracy. Now, with the spread of leisure to all classes and the emergence of what Riesman has called other-directedness, sociability has gained increased importance, certainly in the middle class and perhaps even in the lower class. At the same time, there has been a historical change in America from a view of life emphasizing free enterprise and individual initiative to one emphasizing *social and personal security*. Specifically, what does this mean for sociability?

Consider first the kind of integration with society that is indicated. An upper-class elite devotes itself first to the perpetuation of an elite culture and second to the perpetuation of a society of which it is the privileged upper crust. Sociability is likely to range widely through literature, philosophy, and news of the day. Individuals have both the time and the motivation to keep themselves informed about matters of interest to the elite. But now consider the situation of hard-working persons in the middle range of society.

These people have little sense of special responsibility for the development or transmission of culture. Neither do they have much sense of being responsible for the destiny of society or of their nation. Here we must recognize historical as well as class differences. In the past, and even as recently as the 1930's, problems of national policy seemed both urgent and soluble. Individuals from all walks of life could be concerned and informed. Today, there is a tendency to feel that problems of national policy are so far beyond our power and our understanding that there is no point in even talking about them. The main thing we ask of society is that it let us alone, give us an underpinning of social security, and give us an opportunity to live comfortably in our own private worlds. In short, both social and historical differences serve to place the accent on retreat from society, rather than on either functional or normative integration with it.

The second criterion utilized in setting up the ideal types was the kind of individual needs which dominate the interaction, considering particularly achievement, unconditional acceptance, and uniqueness. It seems evident that there has been a historical change in America in the direction of stressing the importance of economic and psychological security for all, in stressing *unconditional acceptance* and unconditional conformity as the major virtues, and in discounting the importance of achievement and uniqueness. Even in the middle class where achievement is a major virtue at work, there is enough guilt and anxiety about it so that people will disclaim their achievements when they are among friends. Achievement and uniqueness are both likely to be seen as barriers to friendship, rather than as advantages. People almost feel obliged to assure their friends of love and acceptance and to pretend that achievement and uniqueness do not exist.

With respect to involvement of self, there is an increase today in emphasis upon *negative rather than positive aspects of self*. The historical change of most importance is the one that began with Freud and has led to a culture in which mental illness has become a major national problem and psychotherapy of some kind almost a normal part of the experience of large numbers of Americans. It is difficult to realize the difference that this had made in our view of the individual. Two of the most familiar of the discarded myths include the Victorian image of the pure and virtuous individual, now replaced with the idea that man is governed to a large degree by sexual and aggressive impulses; and the free-enterprising image of man as able to get anything he goes after, now replaced with the image of neurotic man governed by anxieties and compulsions. These changes mean that in all kinds of interaction, including sociability, there has been an increase in the degree to which an individual presents himself as weak or troubled and in need of help, and a decrease in the extent to which he presents himself as the accomplished

master of many arts. The shift down the class ladder works in the same way. It has been said that the theme of the upper class is being, while that of the middle class is doing (1). Problems and ambiguities come up for people who are constantly engaged in doing new things or meeting new challenges, in a way which they do not for people who are content with simply being what they have always been.

The last two criteria used to define styles of interaction draw attention to the difference between reiteration of routine experience, progressive advance on the job, and the dramatization of novelty and special interests. The ability to develop *novel and interesting sociable resources* is associated with social class. It takes time to develop what we call conversation resources. A leisure class has time to read, ponder, travel, and garner a variety of interesting ideas and experiences. For middle-class professionals, there is often not much opportunity to do this except on the job, and direct work-related conversation is often taboo at parties and other social gatherings. This may leave people with nothing to talk about. To the extent that they have been doing interesting things in their leisure time, they can talk about them. In the parties and lunches we have observed, for instance, we have heard talk about movies, TV, national sports, card games, automobiles, and summer camping trips. But if people do not have anything new or interesting to report, the next best thing is to offer the latest information on the children or the job, or to fall back onto safe topics that have become well-established in the group. Thus, the relatively busy lives of people in the professional class, at least, work against the development of novelty and new interests in sociability and toward the kind of routine and reiterative interaction which we have called familial.

A historical factor which works in the same direction is that friends are likely to see each other more often now than they did a hundred years ago. It is much easier to be dramatic and entertaining with a person one sees twice a year than with a person one sees every week, or even every day. To the extent that sociability is confined to a group of people who get together "all the time," it tends to become routinized rather than imaginative.

Sociability as a Defense against Anxiety

Let us now take a more psychological view of sociability, considering it as an activity in which individuals may participate partly to reduce anxiety and enhance self-esteem. From this point of view, institutionalized patterns of social interaction are just as much defense mechanisms as are the more familiar individual ego defenses. Individual problems which cannot be handled by one person alone are often taken up and even solved in cooperation with others.

We can begin with the problem of aggression, or destructiveness. This is a problem which is peculiarly salient in the middle class, where assertion is required and hostility forbidden.

The major social defense against one's own destructive impulses is comparable to the individual mechanism of *projection*. Persons join together to form an ingroup based on the assumptions that they all love each other, are all harmless and well-intentioned, and that whatever badness exists in the world is located outside the group in some part of an alien environment. Notice that there are two parts to this defense. The first is the implicit guarantee that everyone in the group will be liked and will like others: the sociable form of unconditional acceptance. The price for this, of course, is that one must not compete with the others while in a social setting, must not refer to his own achievements outside the group except in a depreciatory way, and must accept with apparent pleasure whatever topic of conversation is introduced—usually being content with the lowest common denominator, so that no one will be excluded.

The other half of this defense is what might be called scapegoating, or in many cases, griping. Certain persons or impersonal forces are set up as alien, unfriendly, and even evil. These may vary from the janitor to the head of one's department to President Nasser. Sociable conversation dwells on the outrageous acts of these outsiders, and thereby implicitly affirms the virtues of the persons present.

It should be clear that the defense against hostility involves all the characteristics we have attributed to the familial style: privacy and retreat, unconditional acceptance of persons present, focus on the problematic, and reiterative conversation.

A second source of anxiety is uncertainty about what to expect from the outside world. In a way, this is a concern about destructiveness, too, insofar as it is attack that one fears. But this is not the only reason for wanting a *clear picture of reality*. There are a variety of ways in which a person's decisions and successes depend upon factors which he cannot control but which he would like to understand. It is commonplace to point out that a complex society like our own is inevitably ambiguous. Reality cannot be fully known by perceiving it, and so it becomes known by definition. A major function of sociability is to provide acceptable explanations for aspects of the external world which are either puzzling or threatening. When the criterion for accuracy of judgment cannot be correspondence with observed facts, it becomes correspondence with social reality, which is to say, with the judgments of one's friends. We may also assume that in any group engaged in creating its own picture of reality, pressure will be exerted to make the definition one

that gives maximum self-esteem and protection to group members. Distortion is not an inevitable consequence, but it is a likely one.

In general, the need for social definitions of reality should serve to maximize what we have called sociable styles of interaction. It calls for normative integration, for culture building, and for the assimilation of new experience and new information. However, there are important qualitative variations in interaction which depend on the relation between social reality and what might be called true reality. Playful fantasy and exaggeration of reality may be classified as sociable. Laborious and literal striving for accuracy often resembles work, although it can also remain sociable. The condition that pushes toward the familial style is a compelling need for self-deception. In a power hierarchy, for instance, persons at lower levels of the hierarchy may express resentment about their position through exaggeration of negative characteristics of the upper levels. Or persons with no real chance for succeeding in getting what they want out of life may build their sociability around assurances of success. The greater the amount of self-deception in these processes, the more the sociability will take on the familial characteristics of retreat from society, unconditional acceptance of the individual for his own sake, and reiteration.

There is a third type of anxiety which requires a social defense, and which we have seen particularly in young college students. This is uncertainty about one's own individual identity and worth. The necessary defense is to obtain *reassurance from others about one's own reality and importance*. There seem to be two kinds of situations which create this type of anxiety. One is the situation of transition in which a person is making a change from one way of organizing his identity to another. This is typified by the college student who is breaking away from one family and preparing to establish another, and who often is discarding one intellectual or ideological tradition in favor of another. The other situation which creates identity problems occurs when a person spends the majority of his time and energy in roles which contribute little to his sense of identity, while the roles or activities which are of major importance to him get little attention. This situation can occur for the career-woman-turned-housewife, if she wants to maintain her professional identity and is unable to do so; or for the young girl who wants to get married but must mark time for a while with an uninteresting job and unpromising dates.

The social processes which we have observed and which seem to serve as a reassurance about identity are of two kinds. One is a kind of interminable presentation of the details of one's daily activities. This was observed particularly among the young college girls serving on the staff of a summer resort. Presumably, the unconscious purpose of the individual is to gain reassurance about his own reality by making every detail of his life real to others. If a person has more specialized anxiety about whether a desired characteristic

is really a part of his identity, the interminable presentation may be limited to this one area, rather than covering the full range of daily activities.

The second kind of defense was observed in a professional school, among graduate students in the process of making the transition from a lower middle-class background to professional status. The function of their luncheon sociability seemed to be to recreate the world they were leaving behind them, thereby affirming the continuity between their past identity and their present professional concerns. This is a way of permitting change to occur by maintaining a sociable island where change does not occur, or at least where any necessary change is jointly undertaken with others who have the same past history as oneself.

The interminable presentation places the emphasis on the familial treatment of self, where all reports on difficulties and routine concerns are accepted without question as interesting and relevant; and the protection of a past identity places emphasis on three other aspects of the familial style—retreat, unconditional acceptance, and reiteration.

So far we have mentioned four social defense mechanisms—projection, self-deception, interminable presentation, and the protection of special identities. The fact that all of these are more closely associated with the familial style of interaction than with other styles poses a problem. Must we conclude that the familial style is necessarily or even primarily defensive? For the sake of argument, the answer is a qualified "yes." To the extent that individuals need social support in mastering their own anxieties, we believe they will look for it first in situations where they are offered love and acceptance together with a retreat from the demands of society, that is, in situations characterized by the familial style of interaction. Let us remember that these situations need not be actual families, but may be peer groups or even work groups where familial styles have become established. If we imagine individuals with no need for social support in dealing with anxiety, then we can also imagine familial interaction which is under no pressure to become defensive. However, we prefer to think of social defenses, like ego defenses, as being necessary for all individuals. The difference between pathology and health does not depend on whether or not defenses exist, but on the quality of the defense. The unanswered question about the familial style of interaction, then, is whether it can perform the positive functions of warding off anxiety and maintaining individual self-esteem without introducing the more negative qualities of distortion and false reassurance.

Structural Properties of Sociability

We have already mentioned one of the major properties of sociability which might be called structural, namely the property of *contrast*. Most types of sociability are arranged to provide some specific contrast. Sociability on

the job, for instance, usually contrasts sharply with work interaction. It seems to be given character by the necessity for being antisegmental, antiachievement, and antiproduktive more than by any specific aspirations. In fact, the very notion that one should have aspirations for on-the-job sociability is often rejected with the assertion that in sociability persons should be accepted for themselves and not for their roles or achievements.

The social party offers a different kind of contrast. For one thing, a party is used to punctuate both the weekly and annual calendar. Parties are held on Saturday night, around holidays and birthdays, and to mark turning points in one's occupational or residential career. In this sense, a party may be considered as part of the ritual which permits tight budgeting of time and also as part of the ritual which goes with mobility, whether this mobility be geographical, occupational, or social.

More important, however, is the requirement that a party be festive, fun, and somehow in contrast to the routine of a working day. This means, first, that if there are qualities of work activities and relationships which are unpleasant, these are avoided at a party. It is compulsive avoidance of anything resembling work which accentuates familial rather than sociable qualities. For example, people who spend working hours trying to meet the demands of others, whether in taking care of children or in competitive achievement, may seek a sociable pattern in which they can be passive, relaxed, and not under pressure even to be interesting. Persons who experience uncomfortable isolation or alienation in their daily lives may try to create a party style based on affiliation, with repeated affirmation of uncritical warmth and personal interest. Similarly, people who spend their lives in organizations that are trying to find solutions for weighty social problems may get most enjoyment from parties which are completely frivolous and unconcerned about integration with society.

Now frivolity, of course, is an important sociable quality. The desire for contrast can be one of the major forces leading to the sociable style of interaction. It gives priority to festivity over boredom, to fantasy and frivolity over literalness, to novelty, and to exploration of the dramatic and unfamiliar aspects of a person's identity rather than reiteration of the familiar. When contrast is translated into the search for something interesting, or fun, or playful, it tends to generate sociable qualities. On the other hand, when it takes the form of avoiding unpleasant workaday qualities, it tends to be constricting and to generate interaction whose major purpose is reassurance.

There is one other structural property of party sociability which adds to the resemblance between party groups and family groups. The *guest list* serves to draw a circle around the in-group of people who have been invited to the party, distinguishing them from the vast group of outsiders who have

not been invited. The smaller the party, the more compelling is this implication that the guest list defines a primary group. Note here that the movement of sociability to small living rooms and away from large drawing rooms requires that parties be small. At the small party, there is some implication that people should set aside critical standards, should not pass judgment on themselves or each other, should relax and enjoy themselves. Also, a major source of enjoyment at the party often turns out to be the elaboration of common meanings which are shared by people at the party, but which are unknown to outsiders. This can lead into various kinds of culture building, group fantasizing, and repetitive use of symbols which give identity to the party members as a group. Like contrast, this structural property of a small party as a primary group may get defined so as to stimulate either familial styles or sociable styles. It is not likely to generate work styles, although if the group is united by work ties, work-related symbols and ideas are likely to be an important source of common culture.

The Familial Style in Sociability

We can see that there are a number of forces operating to give sociability familial rather than sociable qualities. This raises the question of whether there is a tendency for people to divide their attention between work-oriented interaction and familial interaction without giving sufficient attention to the possibility of developing sociable styles.

In this connection, note the fifth criterion used for distinguishing between the three styles of interaction, namely, the source of topics to be discussed, or what we have called the conversational resources. Work-oriented interaction focuses on the job to be done, and familial interaction is concerned with routine events and concerns of daily life. No special effort is necessary to keep these resources alive. Sociable interaction, however, does require supplementary effort by participants. If they are to have something interesting to talk about, they must spend some time between sociable events in activities which other people will find interesting.

Right now, there may be a problem in some middle-class groups of knowing what resources to accept as legitimate raw material for sociability. It is easier to specify resources to avoid than resources to accept. For instance, there is some rejection of the older aristocratic culture of literature and philosophy. The newer mass culture of movies, TV, automobiles, and suburban living offers some grist for the mill, but again, people may have reservations about confining sociable conversation to such commercialized topics. Similarly, there is some tendency to talk about work, but such conversation will usually separate men from women, and for this and other reasons, will be defined as unsociable. Perhaps the most acceptable resources grow out of special interests,

leisure-time activities, and areas of personal expertness. But these are the resources that need cultivation, and cultivation, in turn, requires a certain amount of outside time, energy, and even inventiveness. It is our impression that in some of the sociability we have observed, familial styles of interaction arise by default and not by design. It is not that familial styles are preferred, but that the resources and expectations which would be necessary for sociable styles are lacking, or if present, are not recognized.

Manuscript received: February 17, 1958

Revised manuscript received: May 5, 1958

Jeanne Watson
Family Study Center
University of Chicago
Chicago 37, Illinois

REFERENCE

1. McArthur, C., "Personality Differences between Middle and Upper Classes," *Journal of Abnormal and Social Psychology*, 1955, 50, 247-254.

Rating Response Set as a Function of Objective Status Criteria¹

SALOMON RETTIG, FRANK N. JACOBSON, LEO DESPRES, AND
BENJAMIN PASAMANICK, *The Columbus Psychiatric Institute and Hospital*

In a comparative analysis of status overestimators and status underestimators among professionals (12, 13), the authors were confronted with the problem of response set. It was found that persons who expected a significantly higher status from various reference groups than they actually received from these reference groups also rated all other variables higher than individuals who underestimated their own status. It seemed that this overrating tendency was partially due to a response set rather than to their estimation of their status. In developing a method designed to control this tendency of consistently over or underrating all variables, irrespective of the type of variable, questions arose as to whether or not the response set of various individuals could be accounted for in terms of various objective status characteristics such as profession, sex, religion, race, and income.

The methodological implications of response set have been recognized for some time. Cronbach (9) introduced the concept as presenting a serious problem with respect to test validity. The same phenomenon has also been referred to in the literature as *rating habit*, and *response bias*. While response set has presented methodological problems to those who are aware of it and who are working with various types of psychological tests, researchers in related fields such as sociology and social psychology have not given great attention to the problem.

More recent research has given response set wider theoretical implications. Rubin-Rabson (16) reported an inverse relationship between the frequency of noncommittal responses on a conservative liberal attitude scale and the self-sufficiency score on the *Bernreuter Personality Inventory*. Brim and Hoff (8) found a relationship between experimentally induced frustration by means of the Gottschaldt hidden figures and extreme probability responses on a *Desire-for-certainty test*. The extreme probability responses were previously shown to correlate with extreme responses on attitude scales. In an analysis of response bias in an unstructured questionnaire, Berg and Rapaport's data (6, p. 3) revealed significant response sets of undergraduate students in choosing heads over tails on a coin, in choosing the number 3 in a choice of 1, 2, 3, 4 and similar sets to an imaginary questionnaire. Females

¹ This study is part of an investigation of status and job satisfaction among professional people in the State of Ohio. The authors wish to express their appreciation to Dr. I. A. Berg for his critical review of this study.

tended to show a greater frequency of such response sets than males. In an earlier study Berg and Collier (5) found that individuals who were defined as anxious as a result of performance on the *Taylor Anxiety Scale*, chose significantly more extreme responses, i.e., *like much* or *dislike much*, on a test of 60 abstract designs, *The Perceptual Reaction Test*, than persons defined as normal. White females and Negro males selected extreme responses on the *Perceptual Reaction Test* more often than white males.

In a more recent paper, Cronbach suggested that response set was related to external criteria. He stated: "... evidently groups differentiated on external criteria also differ in response sets" (10, p. 15). The present investigators conjectured that those external characteristics which are indicative of objective status, and which are crucial in competition for status, may also be indicative of rating response set. This conjecture derives from the assumption that crucial status characteristics, such as the age, sex, and race of professional persons will make for differential status security among them. This proposition is also suggested by previous findings of the authors in which they demonstrated an inverse relationship between sex, race, and income of professional people and the degree to which these professionals overestimated their own status (12, 13). The authors were of the opinion that the same objective status variables which make for significant differences in status estimation, a behavior *critically* related to objective status, would also make for significant differences in response sets, behaviors *not critically* related to objective status. This hypothesis is most aptly expressed by Berg, who stated:

"Deviant response patterns tend to be general; hence those deviant behavior patterns which are significant for abnormality (atypicalness) and thus regarded as symptoms (earmarks or signs) are associated with other deviant response patterns which are in noncritical areas of behavior and which are not regarded as symptoms of personality aberration (nor as indicators, signs, earmarks)" (7, p. 159).

Should the conjecture be substantiated, the implication is that differences in status security of professional persons can be predicted from deviant responses on *any* set of items as readily as they could be predicted from items that have been carefully constructed to elicit such differences. Or, as Berg has stated: "Stimulus patterns of any type and of any sense modality may be used to elicit deviant response patterns; thus particular stimulus content is unimportant for measuring behaviors in terms of the Deviation Hypothesis" (7, p. 160).

At this point, perhaps, it might be desirable to distinguish between *relative* and *absolute* response sets. The response set of an individual is relative when his responses deviate significantly from the performance of a specific group of which he is a member. Absolute response set represents a significant

response deviation from a criterion which is independent of the performance of a specific group, e.g. chance expectancy. The latter may be characteristic of an entire group or category of individuals.

Absolute response sets, that is, significant response deviations from chance expectancy by the majority of the persons in a group, have been demonstrated by Berg and Rapaport (6). While such sets are curious phenomena, they at most demonstrate cultural biases which render assumptions of chance expectancies doubtful. Absolute response sets are therefore primarily of methodological, statistical concern. They perhaps might also be fruitfully employed in studies of cultural differences.

Relative response sets, that is, significant response deviations from mean group performance, have been studied by Berg and Collier (5) and Barnes (1, 2, 3) for diagnostic purposes. These investigators have shown a significant relationship between the type of relative response set and clinical diagnoses of neurosis, psychosis, and normality. Their findings indicate that relative response sets to apparently neutral items are as predictive of clinical categories as the MMPI which has been carefully constructed for this purpose. From these and other studies it becomes apparent that relative response sets may be fruitfully employed in the area of personality and social psychology, especially when the focus of interest is upon deviations from normative behavior.

The purpose of this paper is to test the hypothesis that there is an inverse relationship between relative rating response set and crucial objective status characteristics such as profession, sex, race, income, age, and religion, among professional persons. Relative response set is defined as a significant tendency of an individual to respond in a manner which appears to be irrelevant to the purpose of the test or interview, and which deviates significantly from the mean performance of the entire sample.

METHOD

The data are based on standardized individual interviews with 400 subjects, representing five different professions, each divided into state employed and nonstate employed. The professions represented were psychiatry, psychology, social work, teaching, and nursing. State employees were selected from six different state hospitals and schools; nonstate professionals came from 27 different clinics, agencies, general hospitals, public schools, and from private practice. There were 40 subjects in each subsample.

During the interviews each subject was asked to rate 22 professional specialties according to their status and prestige. The professional specialties consisted of the ten specialties comprising the sample plus twelve additional professions. The subjects were then asked to rate eight job satisfaction varia-

ables twice, once according to their importance and again according to the degree to which each was supplied to them in their work. All of the above thirty-eight ratings were made using a new technique, referred to as the magnetic board rating technique (4, 14). This technique consists of a thin steel board, on which a scale ranging from 0 to 100 is drawn, and a set of pointed labels, each fastened to a small magnet. The labels, each carrying the title of a profession or job satisfaction variable, are pointed on both ends so that they can be placed on either side of the scale. Finally, the subjects answered a questionnaire pertaining to various indices of objective status (15).

Response Set

The response set of a subject was determined by noting the number of his positive and negative deviations from the mean ratings of the entire sample, on each of the thirty-eight items. The sign test was used to determine if a subject's number of positive or negative deviations exceeded chance expectancy at the five per cent level of significance (two-tailed test). Twenty-six positive or negative deviations out of 38 responses are necessary to reach the five per cent level of significance (11, p. 313). This enabled the investigators to classify each of the 400 subjects into one of three groups. The first group consisted of 136 subjects who were characterized by *overrating* response set. The second group consisted of 219 subjects who did not have any response set. The third group consisted of 45 subjects characterized by *underrating* response set.

RESULTS

Table 1 presents the distribution of over- and underraters according to professional membership group. The distribution does not reach statistical significance ($\chi^2=4.480$, 4 d.f., $p=n.s.$). Psychologists indicate the highest frequency of underraters and lowest frequency of overraters. It appears from these data that professional group membership is not a sig-

TABLE 1
Per Cent of Over- and Underraters, by Professions

	Professions				
	Psychiatry (N=80)	Psychology (N=80)	Social Work (N=80)	Teaching (N=80)	Nursing (N=80)
Overraters (N=136)	.38	.24	.43	.31	.35
Underraters (N=45)	.13	.15	.11	.10	.08

TABLE 2
Sex Distribution According to Response Set Category, in Per Cent

	Overrating (N=136)	None (N=219)	Underrating (N=45)
Male (N=181)	.25	.60	.15
Females (N=219)	.41	.51	.08

nificant factor in response set. Here it must be remembered that the obtained response set was based upon the mean performance of the entire sample of professionals. Had the response set been computed on the basis of mean subgroup performance, that is the mean performance of each profession separately, the results would differ somewhat.

Table 2 presents the distribution of response set and sex. The difference is statistically very significant (chi-square=12.586, 2 d.f., $p < .05$). Over forty per cent of females are overraters as compared to twenty-five per cent of the males. Fifteen per cent of males are underraters, as compared to eight per cent of females. From these data it can be stated that the sex of the professional is significantly related to response set.

Table 3 indicates response set differences in relationship to race. In spite of the small number of Negroes in the total sample (N=26), the racial differences in response set are significant (chi-square=7.634, 2 d.f., $p < .05$). Over fifty per cent of Negroes are overraters, while the majority of whites demonstrate no response set. The proportion of underraters is equal in both races.

Tables 4 and 5 show the proportion of over- and underraters, with both variables, sex and race, controlled. The response set differences between white males and females remain significant (chi-square=8.764, 2 d.f., $p < .05$). For the Negroes, sex differences in response set do not quite reach statistical significance (chi-square=4.268, 2 d.f., $p > .10$). Here it must be taken into account that the number of Negro professionals, especially males, is extremely small. When the proportion of Negro females in each response set category is compared to the proportion of white females, the differences remain sta-

TABLE 3
Race Distribution According to Response Set Category, in Per Cent

	Overrating	None	Underrating
White (N=374)	.32	.57	.11
Negro (N=26)	.58	.31	.11

TABLE 4

Distribution of Whites, According to Sex and Response Set Category, in Per Cent

	Overrating	None	Underrating
Male (N=174)	.25	.60	.15
Female (N=200)	.39	.53	.08

tistically significant ($\chi^2=6.438$, 2 d.f., $p<.05$). From the findings it can be concluded that the variables of race and sex are related to response set, and that these variables operate independently of each other.

Table 6 presents the mean age for each sex and response set category separately, with race controlled. While the age varies consistently with response set category, the differences in age between the overrating response set category and neutral response set category is much larger than the age differences between the neutral category and the underrating category. The age differences between the over- and underrating categories do not reach statistical significance in either group, but the age difference between overrating and neutral response set category does reach significance among the females ($t=2.917$, $p<.05$). It appears that age is a significant variable in response set, primarily among females.

Table 7 presents the mean income for each sex and response set category, controlled for race. The income differences are not in accordance with the hypothesis. The overraters have higher income than both the males and females in the neutral response set categories. The income difference between the overrating and neutral response set category reaches statistical significance among the females ($t=3.276$, $p<.05$) but not among the males. The underraters have the lowest income among the males, but have the highest income among the females. From these results it appears that income is a significant variable in response set, primarily among the female professionals, but that the direction is reversed from that predicted by the hypothesis. It is the overraters who have the higher incomes.

Income and age differences were not computed for the Negro sample because of their small number. The data were also checked for supervisory capacity (number of persons under supervision and supervisory title), religious affilia-

TABLE 5

Distribution of Negroes, According to Sex and Response Set Category, in Per Cent

	Overrating	None	Underrating
Male (N=7)	.29	.43	.29
Female (N=19)	.68	.26	.05

TABLE 6

Mean Age of White Males and Females, According to Response Set Category

	Overrating	None	Underrating
Male	42.6 (44) ¹	40.4 (105)	40.2 (25)
Female	41.0 (77)	35.8 (106)	34.7 (17)

¹ N in parentheses.

tion, number of professional publications, interviewer bias, and for the relationship between the sex of the interviewer and response set. No significant differences were found in any of the above variables.

Discussion

Three independent objective status variables, sex, race, and age were found to be inversely related to response set. A fourth variable, income, appeared to be directly related to response set. Professional group membership, number of professional publications, supervisory capacity, and religious affiliation were found to be unrelated to response set. It cannot therefore be concluded that every aspect of objective status is related to differences in response set.

Further consideration of the objective status variables of sex, race, and age suggest that these variables are qualitatively different from other objective variables. The sex, race, and age of a professional are variables over which the individual has no control, either objectively or subjectively. The individual cannot expect to change either his sex or his race. His age does change, but again the personal aspirations and expectations cannot affect this change. The status derived from these three variables is visible and inflexible.

The status derived from income, professional group membership, religious affiliation, supervisory capacity, or the number of professional publications, is neither visible nor as inflexible. It is most probable that the majority of professionals in this sample expect and actually receive periodic increases in

TABLE 7

Mean Income¹ of White Males and Females, According to Response Set Category

	Overrating	None	Underrating
Male	8.39 (44) ²	7.66 (105)	7.44 (25)
Female	4.71 (77)	3.76 (106)	5.06 (17)

¹ Income in categories.² N in parentheses.

income. It is, indeed, quite probable that aspirations and expectations of increased incomes, more publications, and higher supervisory capacities may operate to offset and compensate for any basic status insecurities derived from sex, race, and age differences.

The variables of sex, race, and age appear to be more crucial to the status and prestige of the professional in our society. They are a more significant source of their status insecurity. Being female, Negro, or of older age puts the professional at a great disadvantage in the competition for professional status. This disadvantage is only partly compensated by increased income, publications, and supervisory capacity at the locus of work. The compensatory effect hardly generalizes to other social or even professional spheres, where the status of the professional is still initially judged by his sex, race, or age rather than by his profession or income.

If the proposition that sex, race, and age are the fundamental sources of the status insecurity of the professional in our society is warranted, then the obtained data support the hypothesis that there is a direct relationship between status insecurity and relative response set. The greater the insecurity, the greater the tendency to overrate variables irrespective of content.

The status insecurity and the anxieties associated with such insecurity may have been precipitated by the interview situation. The highly standardized interviews confronted the subjects with a situation that was novel, and which required numerous complex judgments of unaccustomed precision. Most questions probably elicited very personal reactions during which habitual professional techniques or personal defenses could not be utilized.

Another indication of situational insecurity involves differences in response sets between state employed and nonstate employed subjects. It would be expected that state employed subjects demonstrate greater situational insecurity than nonstate employed subjects since it was known that the study was sponsored by and carried out by professional investigators representing the State Department of Mental Hygiene. The difference in the distribution of state and nonstate employed subjects in the over- and underrating response set categories is statistically very significant ($\chi^2=11.030$, 1 d.f., $p<.05$). Forty-two per cent of the state employed subjects are over-raters as compared to eight per cent of the nonstate employed. When sex and race were controlled, this difference remained significant for the white males ($\chi^2=5.957$, 1 d.f., $p<.05$) but disappeared among the white females. There were no differences in age or professional group membership between the state employed and nonstate employed white male professionals.

Previous experience with tests or interviews, especially as subjects, would be conducive to less situational insecurity. Here, one would expect psychologists to be more confident and have less overrating response set. The findings are in accordance with this suggestion (Table 1).

In considering the findings of this study alongside those of other investigators, a common presupposition becomes apparent. The hypothetical concepts utilized in the explanation of antecedent conditions of relative response sets include references to frustration, anxiety, insecurity, lack of self-sufficiency, and, finally, neurosis and psychosis. These hypothetical conclusions, which come from divergent sources, suggest that response set behavior reflects a state of tension on the part of the individual. The particular form in which the relative response expresses itself may depend, aside from the type of assessment used, upon the intensity of the tension. In any case, the relationships between relative response sets and various forms of tension appears to be worthy of further exploration.

It is, of course, extremely difficult to relate relative response set to other verbal measures of tension or anxiety since the effect of response set upon such other measures would have to be partialled out before their validity could be accepted. On the other hand, the considerable methodological advantage of using response set to measure various forms of tension lies in the fact that such an approach completely conceals the purpose of the measurement.

For the above reasons, the relationship between response sets and various forms of tensions might best be measured experimentally, as has been done by Brim and Hoff (8). An interview situation could be manipulated, for example, in the following ways: varying the place at which the interview occurs, or varying the degree of personal involvement of the respondents by changing the content, the specificity, or the ambiguity of the questions asked. Such an experimental approach might also throw light upon the following problems. What is the nature of the stimuli which elicit response set? How specific, how unfamiliar, and how loaded with affect do the stimuli have to be before they elicit response sets? How aware is the individual of response set behavior and what changes in response set behavior would occur with increased awareness? And finally, what function do response sets serve in the economy and equilibrium of the individual? These and other problems would have to be investigated before the relationship between response set and tension can become more meaningful.

SUMMARY AND CONCLUSIONS

The purpose of this paper is to determine the relationship between response sets and objective criteria of status such as age, sex, race, and religious affiliation. Response set is defined as a significant tendency of an individual to respond in a manner which appears to be irrelevant to the purpose of a test or interview. The data were obtained as part of a larger study of status and job satisfaction among 400 professional people. The subjects were seen

in individual interviews and asked to rate thirty-eight different variables pertaining to status and job satisfaction. Response set was determined by noting the subjects' number of positive and negative deviations from the mean judgment of the entire sample on each of the thirty-eight items. The analysis of the data revealed that sex, age, and race are inversely related to response set in that females, Negroes, and older professionals tended to over-rate. Income tended to be directly related to response set in that overrating white females had a significantly higher income than white females who have no response set. It is suggested that the objective status variables of sex, race, and age are crucial in determining status insecurity among professionals in our society since these variables are visible and beyond the control of the individual. The status insecurity of persons of lower objective status appears to have been precipitated by the highly standardized interview situation during which the subjects had to make judgments of unaccustomed precision.

Manuscript received: January 24, 1958

First revision received: March 19, 1958

Final revision received: July 9, 1958

Salomon Rettig

Research Division

Columbus Psychiatric Institute and Hospital

Columbus 10, Ohio

REFERENCES

1. Barnes, E. H., "The Relationship of Biased Test Responses to Psychopathology," *Journal of Abnormal and Social Psychology*, 1955, 51, 286-290.
2. Barnes, E. H., "Response Bias and the MMPI," *Journal of Consulting Psychology*, 1956, 20, 371-374.
3. Barnes, E. H., "Factors, Response Bias, and the MMPI," *Journal of Consulting Psychology*, 1956, 20, 419-421.
4. Bartlett, C. J., E. Heermann, and S. Rettig, "A Comparison of Six Different Scaling Techniques," *Journal of Social Psychology*, in press.
5. Berg, I. A., and J. S. Collier, "Personality and Group Differences in Extreme Response Sets," *Educational and Psychological Measurement*, 1953, 13, 164-169.
6. Berg, I. A., and G. M. Rapaport, "Response Bias in an Unstructured Questionnaire," *Journal of Psychology*, 1954, 38, 475-481.
7. Berg, I. A., "Deviant Responses and Deviant People: The Formulation of Deviation Hypothesis," *Journal of Counseling Psychology*, 1957, 4, 154-160.
8. Brim, O. G., and D. B. Hoff, "Individual and Situational Differences in Desire for Certainty," *Journal of Abnormal and Social Psychology*, 1957, 54, 225-229.
9. Cronbach, L. J., "Response Sets and Test Validity," *Educational and Psychological Measurement*, 1946, 6, 475-494.
10. Cronbach, L. J., "Further Evidence on Response Sets and Test Designs," *Educational and Psychological Measurement*, 1950, 10, 3-31.
11. Mosteller, F., and R. R. Bush, "Selected Quantitative Techniques," in G. Lindzey, (ed.), *Handbook of Social Psychology*, 1954, Vol. I.

12. Rettig, S., F. N. Jacobson, and B. Pasamanick, "Status Overestimation, Objective Status, and Job Satisfaction among Professionals," *American Sociological Review*, 1958, 23, 75-81.
13. Rettig, S., F. N. Jacobson, and B. Pasamanick, "A Comparative Analysis of Status Overestimators and Status Underestimators among Professionals," in manuscript.
14. Rettig, S., F. N. Jacobson, and B. Pasamanick, "The Magnetic Board Rating Technique," *Journal of Psychology*, 1958, 45, 201-206.
15. Rettig, S., F. N. Jacobson, and B. Pasamanick, "The Measurement of the Objective Status of the Professional," unpublished.
16. Rubin-Rabson, G., "Correlates of the Non-Committal Test-Item Response," *Journal of Clinical Psychology*, 1954, 10, 93-95.

Racial Bias in Peer Ratings of Basic Airmen

JOHN A. COX, *Wright Air Development Center*¹

JOHN D. KRUMBOLTZ, *Michigan State University*

Racial integration has been in effect in Air Force basic training since 1949. Since that time all men have started their Air Force careers by working, studying, and living together with no indication of their ethnic background appearing on any of their official records. The general consensus is that racial integration has been successful in terms of harmonious working relationships and mission accomplishment. Although integration appears to have been successful from a practical point of view, it is still of interest to study the extent of psychological integration that has taken place. To what extent does racial bias still exist in Air Force training flights and how can it be measured?

One indirect approach to this problem consists of analyzing leadership peer ratings by racial groups within Air Force training flights. Are Negroes rated higher by members of their own race than they are by whites? Are whites rated higher by whites than they are by Negroes? If so, does this bias operate in some training flights more than in others? Can Negroes and whites in a given flight agree on the rank order of leadership ability in their flight regardless of the existence of a racial bias? The present investigation is an attempt to answer these questions.

PROCEDURE

The sample consisted of 533 basic airmen in nine flights at Lackland Air Force Base, Texas. Airmen were not assigned to flights at random, but were assigned in the order in which they appeared for processing. In some instances two or more acquaintances may have been assigned to the same flight. The extent to which this occurred is not known precisely. However, these flights were formed according to the standard operating procedures of the Air Force, and the number of friendships within a flight was probably typical of those ordinarily found.

For purposes of this study, only two "racial" designations were utilized:

¹ This report is based on work done under ARDC Project No. 7719, Task No. 17008, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government. The opinions or conclusions expressed herein are those of the authors. They are not to be construed as necessarily reflecting the views or indorsement of the Air Force or of the Air Research and Development Command.

(a) Negro, and (b) non-Negro or white. Although this dichotomy is not ethnologically pure, it has real sociological meaning in the United States. Each of the subjects involved in this study was identified as either Negro or white by a test administrator while the subjects were taking some routine tests during their basic training. To check on the adequacy of the test administrator's judgment, a roster of each training flight was submitted to the Tactical Instructor of that flight for his designation of Negroes and whites. It was found that the test administrator and the Tactical Instructor agreed perfectly on their classifications.

During the late spring of 1955 peer ratings were collected from these men as they were beginning their third week of basic training. These ratings were to be used operationally for the selection of certain flight positions which required leadership ability but no particular technical ability. In each flight of about 60 airmen, each man rated every other man according to the following directions:

Several airmen from your flight will be selected for positions such as squad leader, right guide, assistant flight leader, and so forth. You are asked to rate each member of your flight as to how well you think he can do such a job. The selection of airmen for positions will be made on the basis of your ratings. To make your ratings of the other flight members you will put a number beside each name on the roster as follows:

- 5—Very well qualified for a position, one of the best in the flight.
- 4—A well-qualified man for a position.
- 3—Would not do a bad job or a good one; about average.
- 2—Not well qualified for the job; would do a poor job at this time.
- 1—Poorly qualified for the job; would do a very poor job at this time.

Some airmen were absent from the rating sessions due to reasons such as illness and work details. The absent men were rated by the other flight members but made no ratings themselves. This results in there being a few more people rated than there were people making ratings.

For each man, the ratings received from every Negro in the flight were averaged to form his "Negro peer status" score. Similarly, the ratings from each white person in the flight were averaged to form a man's "white peer status" score. Thus, each individual had two scores. The scores for the Negroes by flight are summarized in Table 1. The corresponding scores for whites are summarized in Table 2.

For both the Negro data and the white data, three hypotheses were tested:

1. Individuals of a given race receive the same mean rating from members of their own race as they do from members of the other race.
2. The mean scores of the flights are equal.
3. There is no interaction effect between flight and the racial source of peer status scores.

TABLE 1
Summary of Raw Data for Negro Airmen
 (N=64)

	Flight										Totals
	1	2	3	4	5	6	7	8	9		
Scores received from Negroes	N	4	4	8	6	4	4	9	13	7	64
	Mean	2.67	4.17	2.68	2.68	2.74	3.00	3.25	2.39	3.49	2.91
	SD*	.47	.43	.94	1.15	.50	.27	.59	.59	.55	.79
Scores received from whites	N	4	4	8	6	9	4	9	13	7	64
	Mean	2.53	3.45	2.56	2.44	2.23	2.92	2.56	2.53	2.90	2.61
	SD*	.59	.32	.58	.49	.44	.22	.28	.51	.60	.53
Scores received from Negroes and whites	N	8	8	16	12	18	8	18	26	14	128
	Mean	2.60	3.81	2.62	2.56	2.48	2.96	2.90	2.46	3.19	2.76
	SD*	.50	.52	.76	.85	.53	.23	.57	.54	.63	.69

* Unbiased estimates

TABLE 2
Summary of Raw Data for White Airmen
 (N=469)

		Flight									Totals
		1	2	3	4	5	6	7	8	9	
Scores received from Negroes	N	51	56	54	54	51	56	50	48	49	469
	Mean	2.49	2.37	2.59	2.40	2.42	2.81	2.67	2.68	2.50	2.55
	SD*	.60	.76	.65	.56	.61	.69	.50	.44	.55	.62
Scores received from whites	N	51	56	54	54	51	56	50	48	49	469
	Mean	2.64	2.81	2.82	2.60	2.53	2.76	2.64	2.92	2.81	2.73
	SD*	.64	.41	.63	.46	.66	.57	.41	.41	.55	.54
Scores received from Negroes and whites	N	102	112	108	108	102	112	100	96	98	938
	Mean	2.56	2.59	2.71	2.50	2.48	2.79	2.66	2.80	2.65	2.64
	SD*	.62	.65	.65	.52	.63	.63	.46	.44	.57	.59

* Unbiased estimates

A two-way analysis of variance was carried out to test these hypotheses. The 5 per cent level of significance was chosen for rejection of the null hypotheses.

To test the assumption of homogeneity of variance within cells, the L_1 test (1, pp. 82-86) was utilized because of the small frequencies within cells. The assumption was satisfied for the Negro data ($p > .05$), but the assumption was not satisfied for the white data ($p < .05$). The effect of heterogeneity of variance on the F distribution has been investigated empirically by Norton and has been reported in Lindquist (2, pp. 78-90). It was concluded that "... where marked (but not extreme) heterogeneity is expected, it is desirable to allow for the discrepancy by setting a slightly higher "apparent" level of significance for this test than one would otherwise employ . . ." (2, p. 83). From the data reported by Norton (2, p. 82), the present writers estimated that setting the "apparent" level of significance requirement at the .025 level would ensure that the "true" risk of a Type I error would be approximately .05 when the homogeneity assumption was not satisfied.

RESULTS AND DISCUSSION

A summary of the analysis of variance for the Negro men is presented in Table 3. The interaction between race and flight was significant. This inter-

TABLE 3
Analysis of Variance for Negro Data

Source of variance	df	ss	ms	F	Hypothesis
Between flights	8	16.8091	2.1011	3.62	Reject
Between individuals within flights	55	31.8987	.5800		
Total between individuals	63	48.7078	.7731		
Between races	1	2.9282	2.9282	7.68	Reject
Interaction: race by flight	.8	3.0503	.3813	3.64	Reject
Net within individuals	55	5.7653	.1048		
Total within individuals	64	11.7438	.1835		
Total	127	60.4618			

action indicates that what was true in some flights was not true (or not true to the same extent) in all flights. Although in general Negroes were rated higher by Negroes than they were by whites, this was not true in all flights. For example, in Flight 8 the 13 Negroes received higher ratings from whites than they did from each other. The remaining contribution to the interaction

effect was due to the fact that Negroes rated themselves high to a greater extent in some flights than in others.

Since individuals were not randomly assigned to flights and since the sampling was performed by flight, the interaction mean square becomes the proper error term for testing the significance of the "between races" effect. Negroes were rated significantly higher by Negroes than they were by whites. Furthermore, it may be observed that the flight mean scores taken as a whole varied significantly from the grand mean. In some flights the average rating for all men in the flight approached the "well qualified" category, while in

TABLE 4
Analysis of Variance for White Data

Source of variance	df	ss	ms	F	Hypothesis
Between flights	8	10.7828	1.3478	2.31	Accept
Between individuals within flights	460	267.9906	.5826		
Total between individuals	468	278.7731	.5957		
Between races	1	7.5009	7.5009	11.92	Reject
Interaction: race by flight	8	5.0363	.6295	8.44	Reject
Net within individuals	460	34.3092	.0746		
Total within individuals	469	46.8464	.0999		
Total	937	325.2000			

other flights the average rating was closer to the "not well qualified" category. This significant difference among flights may be at least partly due to the nonrandom assignment of men to flights.

The analysis of variance of the white individuals' scores is presented in Table 4. Although no significant differences in flight means were observed, the difference between mean scores which whites received from the two races was significant. The whites were rated higher by members of their own race than they were rated by Negroes. The significant interaction was at least partly caused by Flights 6 and 7 where whites received a slightly higher rating from Negroes than they did from whites. In addition, the absolute magnitude of the differences between races was not constant from flight to flight.

To estimate the extent of agreement between races in the rank ordering of leadership, a product-moment correlation coefficient was calculated within each racial group. For the 469 persons in the white sample the Negroes' ratings correlated with the white persons' ratings .77. For the 64 Negroes the two sets of ratings correlated .75. The within-races correlation coefficient was .76. Thus, while each race somewhat favored its own members when making

ratings, the two races were able to agree to a large extent on the order in which members of either race belonged on a leadership continuum.

Additional analysis of the data showed that Negroes rated 21 per cent of the white airmen higher than the mean rating Negroes gave themselves. Also, white airmen rated 36 per cent of the Negroes higher than the mean rating whites gave to whites. Thus, it is apparent from the extent of overlap of the two distributions that whatever bias existed was far from complete.

No attempt was made to compare the actual leadership ability of the two races. If one of the races possessed more "true" leadership ability than the other, such a difference would not affect the interpretation of these results. The primary analysis consisted of comparing the ratings from two different racial groups on one racial group. Since only one racial group was being analyzed at a time, the objective leadership ability of that group was constant. Only the perception of it was different for the two rating groups.

This study suffers from the same factors which hamper the majority of studies of human behavior when data is collected under operating or practical conditions. Random selection of subjects from the population and random assignment of subjects to flights was not possible. The effect of prior acquaintanceship on the results of this study can only be hypothesized at the present time. It seems reasonable to suppose that the majority of such friendships occurred within a given race. Assuming that acquaintances tend to rate each other high, then the more race-bounded friendships, the more apparent racial bias revealed. Thus, if men could have been assigned to flights at random, it seems reasonable to suppose that the amount of racial bias appearing would have been even less than that discovered in this study.

SUMMARY AND CONCLUSIONS

Leadership peer ratings were collected from 64 Negro and 469 white basic airmen in nine flights at Lackland Air Force Base, Texas. For each man two scores were calculated: a "Negro peer status" score based on the mean of ratings received from Negroes, and a "white peer status" score based on the mean of ratings received from whites.

Results of an analysis of variance lead to the following conclusions:

1. On the average, individuals are rated higher by members of their own race than they are by members of the other race.
2. A significant interaction indicates that the first conclusion is not true for every flight or to the same extent in each flight.
3. In general, individuals of both races can agree on the rank order of leadership ability in a flight regardless of race ($r=.76$).
4. The extent of overlap of the two distributions of leadership ratings demonstrates that whatever bias exists is far from complete.

The findings of this study are interpreted as evidence of substantial, though incomplete, psychological integration in the Air Force.

Manuscript received: January 27, 1958

Revised manuscript received: April 28, 1958

John A. Cox, Jr.

Personnel Laboratory

Wright Air Development Center

Box 1557

Lackland AF Base, Texas

REFERENCES

1. Johnson, P. O., *Statistical Methods in Research*, New York: Prentice-Hall, 1949.
2. Lindquist, E. F., *Design and Analysis of Experiments in Psychology and Education*, Boston: Houghton Mifflin, 1953.

Similarity of Personality: A Basis for Interpersonal Attraction?¹

L. RICHARD HOFFMAN, *University of Michigan*

The literature on friendship selection emphasizes the notion that friends are usually more similar to each other than people who are not friends. Richardson (9) has concluded as a result of her summary of studies of mental resemblance among friends that, "Throughout all the traits and range of ages the correlations between the paired scores of friends . . . have been positive with very few exceptions. . . . In traits of temperament, the correlations at most ages have tended to run considerably lower than in intelligence, though they are still positive in a great majority of cases." Cattell (2) has gone even further to say that, "Evidence . . . is overwhelmingly to the effect that in friendship, as we have already seen in marriage, similarity brings congeniality."

These conclusions have been derived from a number of correlational studies which have shown that friendship pairs are more similar to each other than are nonfriendship pairs. In a dissertation by Shapiro (10), it was found empirically that, "Individuals tend to choose as friends persons who are similar to them in single traits and values," and "When considering a profile of objective personality factors, the similarity principle operates for friendship choices." In Shapiro's study, as in all the studies on which Richardson and Cattell have based their conclusions, the friendships were already in existence at the time of the study, so that the causal relations, if such exist, between friendship choice and personality similarity are ambiguous. Despite this ambiguity, in almost every case the authors of these studies imply that similarity of personality precedes and determines the choice of friends. The question nevertheless remains, "Do people select each other as friends because they are similar, or do they become more similar because of the interactions connected with their friendships?"

The similarity hypothesis, however, is not the only explanation which has been offered to explain attraction between people. Especially in group situations, the context of the group has been posited to influence the feelings of the group members toward one another. Although they offer no systematic empirical evidence for their hypotheses, Cartwright and Zander (1) have suggested among other things that where "differences of opinion cannot be reconciled, the cohesiveness of the group will be materially lowered," and that under conditions of "group frustration, some of the participants showed

¹ The author expresses his appreciation to Drs. Theodore Newcomb and Victor Vroom for their constructive criticisms of this article.

evidence of wanting to leave the group when it began to fail" (1, pp. 83-84). These authors are postulating that under certain conditions a lack of interpersonal harmony and experiences of task failure are probably major causes of people's failure to be attracted to the other members of a group. The particular conditions under which different bases for attraction operate are not specified.

In the present paper we shall present data which cast doubt upon similarity as a determiner of attraction, and further data which appear to reinforce and extend the Cartwright and Zander hypotheses.

DESIGN OF THE STUDY

Groups of four people each were assembled in the seven laboratory sections of an undergraduate course in the Psychology of Human Relations for the purpose of studying the effects of group composition on problem-solving performance (6). At the beginning of the spring semester of 1955 and again in the beginning of the fall semester of that year the Guilford-Zimmerman Temperament Survey (GZTS) was administered to all of the approximately 175 students in the class as a measure of their personalities (5). This measure yields a profile of scores on ten relatively independent dimensions of personality: general activity, restraint, ascendance, sociability, emotional stability, objectivity, friendliness, thoughtfulness, personal relations, masculinity. The decision to use this as the measure of personality was based principally on its relatively high reliability (split-half coefficients on the scales ranging from .75 to .87) and the internal validity of the scales resulting from a series of successive factor analyses. In addition, the authors claim that the scale descriptions are derived from "validation information and from clinical experience," principally with the inventories which preceded the development of the GZTS. Although even this literature is meager, studies have been reported in which the GZTS or its predecessors have been shown to be related to other accepted personality inventories (3) and to external criteria of success among clinical psychology students (7) and among executives and supervisors (4). Despite this weak evidence the GZTS appears to be consistently measuring people's perceptions of their own personality traits.

Although even ten relatively unique factors of personality hardly begin to represent the full range of characteristics which determine an individual's personality, we can probably assume that these traits are a sample of a wider range of personality characteristics. Certainly it seems defensible to assume that people who are similar on these ten traits are probably more alike than are people who differ markedly.

Within each laboratory section of approximately 25 students the degree of similarity between the profiles of each pair of students was calculated

using Kendall's tau as the measure of profile correlation (8). For each semester this required the calculation of approximately 2100 taus (a combination of 25 people taken two at a time equals 300, times seven sections).² To determine the reliability of these profile correlations, the scores for a sample of 50 students on odd and even halves of the GZTS were correlated and corrected by the Spearman formula for the shortened length. The median coefficient was .77, and 37 of the 50 coefficients exceeded .64, the value needed for significance at the .05 level of significance.

Homogeneous and nonhomogeneous groups were then formed on the basis of these profile correlations. Homogeneous groups consisted of four people whose taus were significantly positive, indicating that their profiles were more alike than could be expected by chance alone. Nonhomogeneous groups consisted of people whose profiles were unlike, whose correlations were either zero or negative.³ The difference between the mean sums of taus for each type of group was significantly different from zero at the .05 level of significance. These groups worked together throughout the semester as often as possible. Each week the role-playing problems or case discussions compelled the students to interact with the other members of the group and thus get to know each other.

In each semester during the final laboratory session a sociometric questionnaire was administered to all students present. The questionnaire took the following form:

If you were to take this course again and you were told that you were going to be placed in a group of four people for the entire semester, which three (3) people from this laboratory section would you most like to have in your group?

I would most like to have in my group.

I would next most like to have in my group.

I would third most like to have in my group.

The number of choices of people in his group was taken as the measure of the respondent's attraction to his group. When the data for the two semesters

² With the use of the Michigan Digital Automatic Computer (MIDAC) the approximately 2100 correlations required were performed sufficiently quickly to permit establishment of the groups the week following the administration of the Guilford-Zimmerman Survey. The author wishes to thank his wife, Roslyn B. Hoffman, for programming the computation for the computer, and Drs. John W. Carr, III and Cecil C. Craig, representing the Rackham School of Graduate Studies for making the MIDAC available for these computations.

³ Two types of nonhomogeneous groups were originally established: one type in which the correlations between profiles were both high negative and high positive, and a second type in which the correlations were approximately zero. Because the number of groups of the former type was too small for analysis and because the sociometric relationships were similar in the two types of nonhomogeneous groups, they were combined for the present analysis.

were combined there were 70 students in the homogeneous and 128 in the nonhomogeneous groups who responded to the sociometric questionnaire.⁴

The validity of this measure of interpersonal attraction is, of course, difficult to ascertain. Assuming that the respondent read, understood, and responded to the question asked, we can be fairly confident that the people he chose were somewhat more attractive to him than were the other students in the class. We do not know the intensity of this attraction nor can we be sure, despite the apparent similarity, that this measure is equivalent to the measures of friendship choice used in previous studies.

Some rather tenuous evidence for the validity of the measure of sociometric attractiveness was obtained, however. During each semester data on the problem-solving activities of the groups were collected for two problems, each on a different day. Attraction should lead to a desire to participate in the group's activities. We may assume, therefore, that one indication of a person's lack of attraction for his group would be his absence from class. The number of in-group preferences made by the members of groups which were intact for both problems (16 groups)—i.e., where all four members were in class on both days—was compared with the number made by members of groups which were not intact. If we can assume that the two problem days are representative samples of the 'groups' attendance records, then the data in Table 1 indicate that the members of groups which were continuously

TABLE 1

In-Group Preferences by Members of Groups Intact and Not Intact for Two Problems

	Intact † groups	Not intact † groups
Mean *	1.58	1.32
Standard deviation	0.943	0.961
Number of students	62	132

* T-test of the difference between means of the intact and not intact groups is significant at the .05 level of significance (one-tailed test).

† The word "intact" identifies the members of groups where all four members were present for two problems and "not intact" identifies students in groups where one or more members were absent on either of the two days.

in class were more attracted to each other than the members of groups whose members were often absent. The mean in-group preference for members

⁴ Because all assignments to groups of the homogeneous and nonhomogeneous types were done within laboratory sections, a sizeable number of students were left unassigned in each section, since the groups they formed did not meet the criteria established for assignment to a type. Students in this category plus a number of assigned students who were absent constitute the difference between the potential 350 students in the original population and the 198 used in this analysis.

of intact groups is significantly greater (by one-tailed t-test) than that of the members of the nonintact groups. These data indicate that the sociometric has some validity as a measure of attraction to the group.

RESULTS

The mean in-group preferences for the homogeneous and nonhomogeneous groups are presented in Table 2. Although the direction of the difference in

TABLE 2
In-Group Preferences in Homogeneous and Nonhomogeneous Groups

	Homogeneous	Nonhomogeneous
Mean *	1.49	1.36
Standard deviation	1.011	0.993
Number of students	70	128

* T-test of the difference between means of homogeneous and nonhomogeneous groups was not significant at the .05 level of significance.

attractiveness between the two groups is as expected, the .13 difference between the means is readily attributable to chance factors in view of the wide variations found within each group type. Thus the assumption that similarity of personality leads to personal attraction has not been borne out by these data.

Since the similarity hypothesis was not confirmed, other factors were explored to see what was accounting for the variations among groups. In exploring the correlations of other factors with sociometric preference within each type of group two opposing relationships were found which appear to support a situational explanation for attractiveness.

The number of women in the group was unrelated to sociometric choice in the homogeneous groups, but a significantly negative relationship was found in the nonhomogeneous groups. Groups with one or no women were significantly more often above the median attractiveness than groups with two or three women. (Unfortunately the predominantly male composition of the class permitted the formation of only one all-female group, a homogeneous one. We were unable to test the possibility that the negative relationship between group attractiveness and the number of women in the group is non-linear.) In other words, personality differences and sex differences seemed to interact to produce an unattractive group atmosphere. In the homogeneous groups the differences in sex composition of the groups appeared to make no difference in their attractiveness for their members (see Table 3).

On the other hand, in the homogeneous groups the group's problem-solving success was directly related to the group's attractiveness, whereas it was

TABLE 3

Relationship between Number of Women in the Group and In-Group Preference

Homogeneous groups (N=18 groups)		
Number of women in group	Below median * in-group preference	Above median in-group preference
0 or 1	7	6
2-4	3	2
Chi-square is not significant		
Nonhomogeneous groups (N=25 groups)		
Number of women in group	Below median in-group preference	Above median in-group preference
0 or 1	10	4
2 or 3	2	9
Chi-square is significant at .05 level		

* Median in-group preference was 1.35 for all groups.

unrelated in the nonhomogeneous groups. The rank-order correlation for the 13 homogeneous groups which were wholly intact for the Mined Road Problem between the quality point score on that problem and mean in-group preference was .55, which is significantly different from zero at the .05 level of significance.⁵ (This problem requires the group to develop a method for crossing a mined road using materials provided and yields a range of scores based on a content analysis of the solutions. It was used for this analysis because it had been solved the week prior to the administration of the sociometric, and it was felt that the group's performance on that problem would be most salient for the members.) In the nonhomogeneous groups, on the other hand, no relationship ($Rho = -.02$) was found between the scores on the problem and the mean in-group preferences of the groups.

Thus the group's attractiveness was related to successful goal achievement but not to the sex composition of the groups of the homogeneous type, and to the sex composition but not to goal achievement in nonhomogeneous groups.

DISCUSSION

These results suggest that the relationship of personality similarity to interpersonal attraction is a complex one, and not a monotonically increasing one

⁵ Rank-order correlations between the mean in-group preference and scores on the Mined Road Problem were used instead of product-moment correlations to avoid the restriction in range on the latter variable caused by the superior problem-solving performances of the nonhomogeneous over the homogeneous groups (6).

as suggested by Richardson and Cattell. Heterogeneity, either of personality or of sex alone, appears to have no direct effect on the attraction of group members to each other. Any differences arising from such heterogeneity are probably fairly easily reconciled without harming the relations in the group. When, however, sex heterogeneity is combined with personality heterogeneity, a sharp drop in attraction occurs. The data suggest that there may be some point beyond which differences among the group members are so great as to produce destructive interpersonal relations in a group. In the present study only sex differences were combined with personality differences to create unpleasant groups. Perhaps ability or attitudinal differences in combination with personality differences would have the same effects. Excessive heterogeneity of group structure appears to create a condition under which the Cartwright-Zander hypothesis concerning the negative effects of interpersonal differences will hold. Thus, although the hypothesis that personality similarity leads to interpersonal attraction was not confirmed, the commonly found moderately positive correlation may be accounted for by the presence in the group studied of a minority of people who are extremely different from each other and who fail to be mutually attractive.

The data suggest also an interesting way of specifying the conditions under which different bases of attraction will be used by group members. The hypothesis offered here is that a group will be attractive to its members to the extent that they are able to solve their most difficult problem. Actually on the basis of the data the hypothesis should be stated negatively, that a group will be unattractive to its members to the extent that they fail to solve their most difficult problem.

This hypothesis derives from a consideration of the contrasting correlations with in-group attractiveness found in the homogeneous and nonhomogeneous groups. It appears that the structure of these groups produced different types of problems by virtue of their varying personality mixture. Homogeneity of personality restricted problem-solving effectiveness, whereas heterogeneity created the potential for interpersonal differences.

Thus the least attractive groups of the homogeneous type were those least effective on the Mined Road Problem. Since the homogeneous groups generally produced poor solutions to this problem, the solutions of the unattractive groups were grossly and probably obviously inadequate. In contrast, the least attractive of the nonhomogeneous groups were those in which sex differences were combined with personality differences. We have inferred earlier that this combination of differences probably created an irreconcilable conflict which contributed to the lack of attraction in these groups. Conversely, where the personality differences were uncomplicated by sex differences the group

members showed a higher than average in-group attraction, suggesting that if differences arose they were reconciled to most people's satisfaction. For each type, the members were least attracted to the others in the groups which failed to solve the major problem which the group structure created. Furthermore, the factor (problem-solving success) which distinguished attractive from unattractive groups of the homogeneous type, was unrelated to in-group attraction in the nonhomogeneous groups, and vice versa.

SUMMARY

Groups of students were initially established in an undergraduate psychology course according to the interrelations among certain measures of their personalities; some groups (homogeneous) with similar personalities, and other (nonhomogeneous) with dissimilar personalities. The hypothesis that similarity of personality leads to personal attraction was examined and found wanting. No greater preferences were found among the homogeneous groups than among the nonhomogeneous groups.

Correlations were found between sociometric choice and problem-solving performance in homogeneous groups and between sociometric choice and sex composition in nonhomogeneous groups. The correlation of sociometric choice with sex composition in homogeneous groups and with task achievement in nonhomogeneous groups was zero. On the basis of these data the hypothesis was suggested that attraction to other group members is a function of the success of a group in accomplishing the more difficult task its personality structure creates, whether this be problem solving or group maintenance.

Manuscript received: January 17, 1958

Revised manuscript received: July 3, 1958

L. Richard Hoffman
Department of Psychology
University of Michigan
Ann Arbor, Michigan

REFERENCES

1. Cartwright, D., and A. Zander, *Group Dynamics*, Evanston, Ill.: Row, Peterson, 1956.
2. Cattell, R. B., *Personality*, New York: McGraw-Hill, 1950.
3. Gilbert, C., "The Guilford-Zimmerman Temperament Survey and Certain Related Personality Tests," *Journal of Applied Psychology*, 1950, 34, 394-396.
4. Guilford, J. P., "Temperament Traits of Executives and Supervisors Measured by the Guilford Personality Inventories," *Journal of Applied Psychology*, 1952, 36, 228-233.
5. Guilford, J. P., and W. S. Zimmerman, *The Guilford-Zimmerman Temperament Survey*, Beverly Hills, Calif.: Sheridan Supply Co., 1949.

6. Hoffman, L. R., "Homogeneity of Member Personality and Its Effect on Group Problem-solving," *Journal of Abnormal and Social Psychology*. In press.
7. Kelly, E. L., and D. W. Fiske, *The Prediction of Performance in Clinical Psychology*, Ann Arbor, Mich.: University of Michigan Press, 1951.
8. Kendall, M. G., *Rank Correlation Methods*, London: Griffin, 1948.
9. Richardson, H. M., "Studies of Mental Resemblance between Husbands and Wives and between Friends," *Psychological Bulletin*, 1939, 36, 104-120.
10. Shapiro, D., "Psychological Factors in Friendship Choice and Rejection," Ph.D. dissertation, University of Michigan, 1953.

The Effect of Talkativeness on Ability to Influence Group Solutions of Problems¹

HENRY W. RIECKEN, *University of Minnesota*

On the basis of his studies of five-man groups discussing human relations problems and deciding on solutions for them, Bales (1) has reported a clear tendency for the member who does the most talking to be credited by his fellow-members with having contributed most to the solution of the problem. Other investigators (2, 5) have obtained substantially the same result in different situations with different problems.

Several points about Bales' finding are worth noting. First, the obtained relationship is between quality of ideas and sheer quantity of interactive behavior (almost entirely verbal behavior). Second, the relationship is established by subjects' ratings of quality rather than by any objective assessment of the worth of the ideas; and, it is the subjects' perceptions of which member contributed the best ideas rather than some objective procedure for identifying their source. It could readily be argued that the active participants are neither the best possible judges of the intrinsic merit of an idea expressed during a meeting, nor the most accurate observers of who said what during a discussion. Finally, it is clear that the problems to be solved by the subjects in the several studies cited above did not have uniquely "best" solutions, and hence, that it might be hard to make dependable objective judgments of the quality of any person's contributions.

These ambiguities make it difficult to interpret Bales' finding. In particular, one would like to know whether the relationship is an artifact of the procedure for assessing quality and source of contributions; whether, perhaps, it is found only when there is no unique best solution for the problem; or whether it is indeed true that the best ideas for solving the problem come from the member who does the greatest amount of talking.

One way to answer these questions is to arrange a problem-solving discussion in which the most frequent interactor in the group (or the least frequent interactor in the alternative treatment) is equipped with the best solution to a problem, and then to observe whether groups in which great talkers have the best solution are more likely to agree upon that solution than are groups in which infrequent interactors have it. The experiment reported below was designed to fit this purpose.

¹ This research was carried out under contract with the Office of Naval Research (Contract number N8-ONR-66216). The assistance of James Banks in performing the interaction scoring and analyzing the questionnaire data is gratefully acknowledged.

PROCEDURE

Groups of four members were formed from voluntary subjects enrolled in elementary psychology courses. In general, subjects were previously unacquainted with each other, although no special effort was made to ensure that the members of any group were complete strangers.

Instructions to Subjects

Subjects were told that the experimenter was interested in how groups solved "human relations problems," and that they would be given several such problems. Each member would be given a written description of the problem and a chance to read it. Then the group would discuss the problem and try to reach a solution.

Subjects were further told that each of these problems had actually occurred in some industry or business and had been successfully solved, so a proven solution was known. They were instructed that when they had reached a solution that all or a majority could agree upon, they were to write it out on a sheet of paper.

Further instructions explained that E would observe the discussion from behind a one-way screen and set the time limits—3 minutes of silence for reading the problem-statement, and 25 minutes for discussion, with a warning signal 5 minutes before the end of the discussion period. Finally, E informed the subjects that he would distribute a questionnaire at the end of each problem discussion in order to secure some additional data.

Discussion Problems

Three human relations problems were presented to the subjects for discussion, the first two of which were used simply to develop an ordering of quantity of interaction among group members. The first two problems had the same general form as the third, namely: a situation in a business or industry in which the behavior of an employee was interfering with the task of the organization.

The third problem was superficially similar to the first two but actually did have a uniquely best solution that was not likely to be discovered by student S's. Furthermore, the essential nature of this solution depended on an insight that could be communicated easily and quickly to an individual. By secretly giving one member of the group the "insight" solution, the experimenter could be sure that the "best idea" would be initially suggested by only one person in each group and thus the *actual* (rather than the perceived) source of "the best idea" could be unequivocally determined.

This third problem was an adaptation of Maier's (4) "parasol assembly

problem." It concerns an assembly line on which one worker is substantially slower than his fellows so that pieces pile up at his station while those behind him are idle or working at less than full speed. The slow man suffers from lack of manual dexterity but, for various reasons, it seems undesirable to discharge him. The question put to the subjects is: What should be done to increase the production of the line?

Hint. All four subjects in each experimental group were given a detailed, two-page statement of the foregoing problem at the start of the third discussion period. One member of each group, however, found the following additional matter typed at the bottom of the second page of his problem description:

You are the only one in this group to whom the hint below is given. You may use the information in any way you wish but *you must not reveal that you have been given special information.*

HINT: The best solution to the problem is to have the men exchange places periodically, progressing from one position to the next in the direction of the flow of work. In this manner, the fast workers would reduce work piled up in positions occupied by slow workers, thus making production dependent upon the ability of the average man rather than on the ability of the slowest worker.

The hint contains compactly all the essential elements of what Maier calls the "elegant solution" for the parasol assembly problem. Presumably, since he was equipped by the hint with the requisite insight, a reasonably intelligent subject should be able to present a dramatically novel solution for the problem, and be able to argue convincingly for his proposal.

Measurement Techniques

As mentioned above, the first two problems for discussion provided an estimate of the amount of talking each member of a group usually did in that group, so that E could decide which member was to receive the hint. A simple measure of total elapsed time of speaking was used to determine interaction rank since it gives the same ordering of group members in groups of the size used in this experiment as does Bales' more complicated method.

The total time talked on the first two problems was used to decide which member of the group would receive the hint on the third problem. In half the groups the hint was given to the member who had done the greatest amount of talking (the "High" treatment); in the remaining groups, it went to the least frequent interactor (the "Low" treatment). A chance procedure was used for deciding the treatment for each group. There were a few groups in which the criteria could not be met. In two groups, one man out-talked other members on the first problem but was himself out-talked on the second. In one group one man talked least on the second problem but next-to-least on the

first. In these three groups the strict criteria were violated and the hint assigned to the man whose *total* interaction on *both* problems was greatest (or least) when compared with his fellow members'.

A second observer joined the timing observer behind the one-way screen for the third problem discussion. The second observer's job was to keep a record of the occasions on which the holder of the hint mentioned any of the essential facts or arguments of the elegant solution. The second observer recorded the time, the item mentioned, and the response that *other* members of the groups made to the hint-holder's statement. Responses were classified as *supportive* (+) when they indicated agreement, interest, or desire for elaboration; or as *rejective* (-) when they showed disagreement, disbelief, or consisted of counterarguments.

Questionnaire

The questionnaire distributed at the end of each problem provided data about the participants' perception of each other and their activities. Each subject ranked his fellow members on: "Who contributed most to the solution of the problem?"; and "Who did the greatest amount of talking?" Another question asked whether there was "any single idea or suggestion that formed the basis for the solution which the group chose," and, if there was, to identify the contributor of the suggestion and to state briefly what it was. Finally, one question concerned the respondent's satisfaction with the solution achieved by the group.

OBJECTIVES OF THE EXPERIMENT

The specific questions the present experiment was designed to answer are three.

First, will the group member who characteristically does the greatest amount of talking be perceived as having contributed the most to the solution of the problem?

Second, is the "top man's" contribution to the solution a result simply of his position in the interaction hierarchy, or is it due to the superiority of his information, opinions, or suggestions as such? If a "bottom man" is equipped with superior information—the hint—will he be as able as a top man to get the elegant solution accepted?

Third, how accurate are group members in judging who contributed most to a solution? Do they more readily recognize a contribution if it comes from a top man than from a bottom man?

Finally, what insight can be gained as to the processes that help a top man or hinder a bottom man in contributing to group problem-solving, if indeed that turns out to be the case? This question is exploratory in nature, and

we shall be less concerned with the statistical significance of findings than with their power to suggest explanations for what we observe.

RESULTS

The first question is readily answered and the answer confirms the findings of earlier studies. Figure 1 reproduces the graph Bales (1) used to present his data and, alongside, the data from the present experiment. It is clear that the higher the individual ranks in amount of interaction initiated, the higher he is ranked in terms of his contribution to the solution of the problem. The results of the present experiment exhibit an even clearer relationship between these variables than Bales found, although the increased sharpness of the differences between men may be attributable in part to the smaller size of groups used in the Minnesota experiment.

Table 1 provides the data relevant to the second question. The top man is more often able to get the elegant solution accepted by his fellows than is the bottom man, even though the latter is, by design, equally well equipped with information and suggestions. When the top man has the insight needed to solve the problem elegantly, the group accepts this solution more than two thirds of the time; when the bottom man has the same information, the elegant solution is rejected in more than two thirds of the groups. This result does not achieve the conventional level of significance (by Fisher's exact test, $p=.08$), but it is consistent with Bales' earlier finding. It seems reasonable to conclude, therefore, that it is probably not the superiority of the top man's information, suggestions, or opinions as such that lead him to be seen as having contributed most to the solution. Rather, his influentiaity seems to be the result of his status as the most frequent talker in the group or of some personal attribute associated with this status.

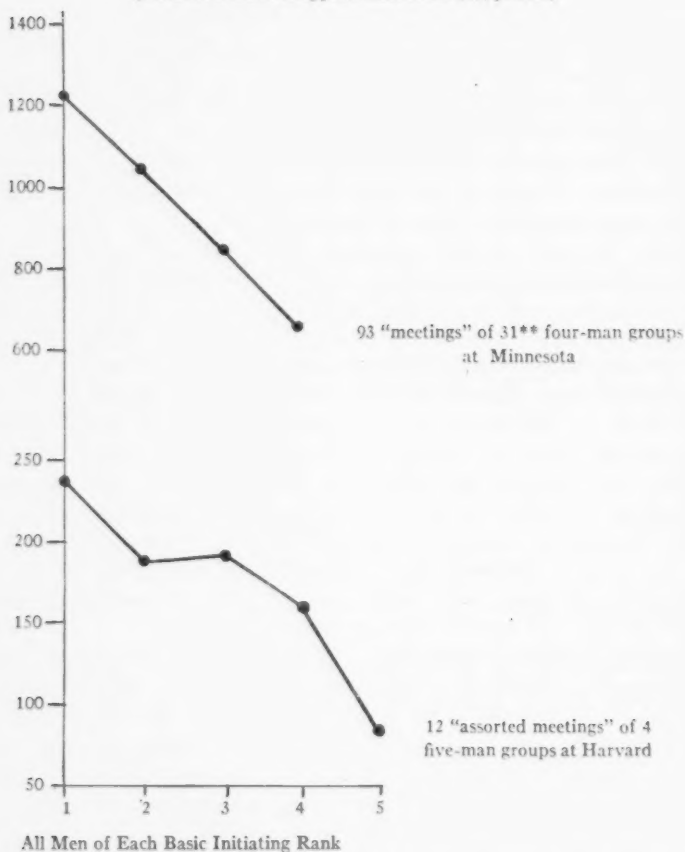
Our third question concerns the accuracy of group members in identifying the source of a contribution to the solution of a problem. Since our attention will be given principally to the relationship between talkativeness and perceived contribution, it is worth while to point out initially that group members are fairly accurate in identifying the individuals who did the greatest and the least amount of talking during a discussion. Combining all groups on all problems, and excluding self-ratings, subjects' rankings of the most talkative member were accurate 68 per cent of the time and their rankings of the least talkative member were correct 72 per cent of the time. (Interestingly enough, subjects did not improve in accuracy from Problem 1 to Problem 3).

Proceeding now to the heart of the question, Figure 1 has already shown there is a distinct tendency to identify quality of contribution with amount of talking. Even when the possible effect of having the hint is eliminated

FIGURE 1

"Total Number of Votes Received" in Answer to Question About Who Contributed Most to the Solution of the Problem.*

(Data from Bales is approximate from (1), p. 146.)



* "Total Number of Votes Received" calculated by summing scores derived from ranks assigned by each member to every member (including self-ranking) on each problem in answer to "Who Contributed Most to the Solution of the Problem?", and pooling scores for all men of each basic initiating rank. Thus, a man receiving rank 1 on the question gets a score of 4; a man receiving a rank of 2 gets a score of 3, etc. This procedure appears to be the one used by Bales to derive his Chart III (p. 146 in (1)).

** Data for one group are incomplete on this question.

TABLE 1

Number of Groups Accepting or Rejecting the "Elegant Solution" by Status of Hint-holder

Acceptance of "elegant solution"	Hint-holder's interaction status	
	High	Low
Accept	11	5
Reject	5	11
Total	16	16

by considering only Problems 1 and 2, it turns out that 59 per cent of all the first-place rankings in answer to the "contribution" question were assigned to the member who had actually done the greatest amount of talking. Table 2 makes this tendency quite clear.

TABLE 2

*Mean Rank * Assigned to Hint-holder in Reply to Question: "Who Contributed Most to the Solution of the Problem?" (Self-ratings Excluded)*

		Problem 1	Problem 2	Problem 3
High	Accept	1.27	1.73	1.27
	Reject	1.80	1.93	2.53
Low	Accept	3.53	3.20	1.80
	Reject	3.73	3.24	3.45

* 1=Contributed most

The data on Problem 3 (in Table 2) suggest strongly that subjects do not simply and completely confound amount of talking and quality of contribution. At the end of Problem 3 other members of the group rate the hint-holder high in amount of contribution when the group accepts his elegant solution, and accurately rate him down when it is rejected. Thus, having the hint tends to boost the ranking that the low-talking hint-holder receives from others on contribution, *provided that he is able to get his proposed solution accepted.*

On the other hand, it is not entirely clear that group members can distinguish between quality and quantity of contribution, for there is some tendency to rate a high-talking hint-holder as having contributed more than a low-talking one even when the group accepts the elegant solution. Further analysis shows that in the High Accept cell, 82 per cent of the subjects correctly attribute the source of the greatest contribution to the hint-holder; whereas, in the Low Accept groups, only 60 per cent make the correct attri-

bution. It may well be that there is a general tendency toward accuracy in recognizing the personal source of a good suggestion, but group members appear to be more accurate, so to speak, in recognizing an identically good contribution when it comes from a more talkative rather than from a less talkative member.

Finally, the replies to the question: Was there any single idea or suggestion that formed the basis for the solution? tend to support the other findings on attribution. As Table 3 shows, in the High Accept groups most members

TABLE 3

Number of Nominations of the Hint-holder in Reply to Question: "Was There Any Single Idea or Suggestion that Formed the Basis for the Solution? If So, Who Contributed the Idea?" (Self-ratings Excluded)

	High accept groups	Low accept groups
Hint-holder nominated	25	7
Other person or no nomination	8	8

(excluding the hint-holder himself, of course) correctly identify some element of the hint and correctly attribute it to the hint-holder. But in the Low Accept groups fewer than half the members recognize the suggestion of the hint-holder and credit him with it; slightly more than half of the subjects in these groups fail to identify the hint, or say there was no single suggestion on which the solution was based, or attribute the hint to someone in the group other than the hint-holder.

These findings tentatively suggest two possible relationships between talkativeness and perceived contribution. Perhaps the participants in a discussion confound quantity with quality in such a way that the greatest talker is also seen as making the best contributions. Such a generalization could hold even if the members of the group made no mistakes in attribution of the source of a particular idea or suggestion. On the other hand, perhaps inherently good ideas are recognized as such but are incorrectly attributed to the wrong person when these ideas are produced by an untalkative member. This latter possibility suggests some perceptual principle of congruence and is almost reminiscent of the findings of Horowitz, Lyons, and Perlmutter (3) regarding the tendency of group members to attribute liked statements made during a meeting to persons whom the attributor likes.

To be sure, neither of these two possibilities is unequivocally demonstrated in the present data, but they suggest that subjects' reports of the quality of contribution to discussion tend to underestimate the contributions of those who characteristically have relatively little to say.

The fourth question to which the present experiment was addressed asked what factors might help to account for the relative success of the top man and the relative failure of the bottom man to get the elegant solution to Problem 3 accepted, and it is to this exploratory question that we now turn.

One rather obvious possibility for explaining differential acceptance invokes the intelligence of the hint-holder: more intelligent men in this position have their suggestions accepted, less intelligent ones fail. This plausible explanation was investigated by comparing the scores of hint-holders and nonhint-holders on a standard test of measured intelligence, the A.C.E.² Since the explanation places its reliance on relative intelligence, the comparison was made by ranking the percentile scores of the four men in each discussion group and comparing the relative ranks of hint-holders in each of the four conditions of the experiment. The distribution of ranks when assessed by the median test yields a chi-square value of 0.282, which has a probability of occurrence by chance of between .95 and .98. Thus there is every reason to believe that there are no significant differences in relative intelligence among the various categories of hint-holders and, in fact, that they are remarkably similar. Differences in intelligence appear to play no part in accounting for the differential acceptance of the elegant solution, nor in the differential talkativeness of the various hint-holders. Finally, there is no evidence of interaction between intelligence and talkativeness in relation to acceptance.

A second possible explanation of differences in acceptance of the elegant solution revolves around the fluency, persuasiveness, and skill of the hint-holder—that is, the extent to which he uses available information, offers convincing arguments, and pushes the elegant solution. Table 4 presents

TABLE 4
Activity of Hint-holder on Problem 3

Condition	Number of elements of hint mentioned at least once		Total number of times any element of hint or argument for elegant solution was mentioned	
	Mean	SD	Mean	SD
High-accept	5.73	1.41	25.18	8.07
High-reject	2.80	1.47	12.80	8.98
Low-accept	4.00	2.28	17.00	9.82
Low-reject	3.73	0.85	19.09	10.16

² The American Council on Education Psychological Examination for College Freshmen is ordinarily administered to entering students at the University of Minnesota and percentile scores based on national norms were available for all the subjects except for one hint-holder in a High Reject group and one in a High Accept group. These two groups had to be dropped out of the analysis of the intelligence test scores and these results are therefore based on an N of 30.

data on the number of elements of the hint that its holder brought out in discussion and the number of times he mentioned some element or offered rebuttals to counterarguments. These findings are inconclusive in that the only statistically significant differences (at the .05 level or better) are between High-Accept and High-Reject hint-holders on both measures; and between High-Accepts and Low-Rejects on number of elements mentioned. The direction of some of these differences, however, is suggestive.

First, it appears that High-Accept hint-holders are more differentiated from others by the number of hint elements they mention than by the total number of mentions they make. In other words, it may be that High-Accepts are using a superior strategy, namely: more varied arguments and a broader range of suggestions and information. In this sense, the High-Accepts are more fluent and skillful persuaders.

On the other hand, there is virtually no difference between the successful and the unsuccessful Low hint-holders on either measure. They are remarkably alike in activity, and fluency or skill of the hint-holder does not seem to distinguish acceptance from rejection in this group.

Finally, the differences in activity between the successful and unsuccessful High talkers engages our attention. Successful high-talking hint-holders mention more than twice as many elements of the hint and speak in favor of it twice as often as their unsuccessful counterparts. The explanation of this curious state of affairs can only be conjectural at this point, but perhaps it is that the unsuccessful high-talkers failed because they personally were not convinced that the hint was an elegant solution and therefore did not push it very hard. Rather, perhaps, they took the stance of an instructed delegate who is not in sympathy with his party's position and thus advocates it weakly—only enough to pay lip service to his role.

Some support for such an interpretation comes from the data on the participants' ratings of their satisfaction with the solution the group reached on Problem 3. Understandably, the successful hint-holders in both High and Low treatments were most satisfied with the solution, their average ratings being 3.73 and 3.60 (on a 4-point scale ranging from 1. "Dissatisfied" to 4. "Quite Satisfied"). Equally understandable is the average rating of 1.91 by unsuccessful hint-holders in the Low condition, since their groups had rejected the elegant solution. But the unsuccessful hint-holders in the High condition returned an average satisfaction rating of 3.00. It seems reasonable to conclude that these subjects were unconvinced of the values of the hint, did not try hard to push it, and were rather relieved when it was rejected. If this explanation is correct, we are forced to add a restriction to the hypothesis that a high-talker can exercise greater influence on the deliberations of a group: he can do so only if he is convinced of the validity of what he advocates. But, under such a restriction, the evidence for a positive rela-

tionship between talking and influence would probably be stronger. If all the high-talkers had been convinced as well as equipped with the elegant solution, there might have been no cases in the High-Reject cell.

If we now turn our attention away from the hint-holder's behavior and look at the responses of the rest of his group to him, we find sharp differences between conditions of the experiment. Table 5 presents data obtained by

TABLE 5

Amount of Support and Rejection Received by the Hint-holder on Problem 3

Condition	Number of supportive remarks made to hint-holder		Number of rejective remarks made to hint-holder	
	Mean	SD	Mean	SD
High-accept	10.55	4.03	14.82	8.22
High-reject	2.60	1.36	11.00	8.25
Low-accept	10.20	2.93	19.00	11.08
Low-reject	2.36	2.10	19.91	11.38

scoring the responses made by other group members to each mention, by the hint-holder, of some element of the hint. It is immediately apparent that the chief difference in response to successful (as compared to unsuccessful) hint-holders is the relatively large amount of support the former receive. All comparisons between successful and unsuccessful hint-holders in amount of support are significant at the .01 level. On the other hand, the amount of opposition, or rejection of the hint-holder's remarks, shows surprisingly little variation among conditions. None of the comparisons between groups of hint-holders on the measure of opposition even approaches statistical significance. In short, success in gaining acceptance for the elegant solution seems to depend not on amount of opposition aroused, but on amount of support the hint-holder musters. The suggestions and opinions of unsuccessful hint-holders are not opposed so much as they are ignored. Perhaps one part of the explanation as to why a highly talkative member is able to influence a group is that his fellows pay attention to him, whereas they ignore an un-talkative person.

Yet this last generalization does not explain the success of hint-holders in the Low-accept condition, for these men apparently do not succeed through becoming more talkative. Among the Low hint-holders as a whole, possession of the crucial information appears to have an inflationary effect on the amount of talking they do: 6 out of 16 of them move up to third place or higher in the interaction hierarchy. But only one of these six is successful in getting the elegant solution accepted. The combination of increased activity and superior information is not a guarantee of success.

If, instead of looking at the activity of the hint-holder in the Low-accept condition, we examine more carefully the responses of his co-members, we find an interesting and suggestive fact. In all five of the Low-accept groups there is at least one member whose response pattern shows more support of than opposition to the hint-holder's arguments for the elegant solution. (Interestingly, in four of these five groups it is the "second man"—to use Bales' phrase—not the top man who takes this role of providing support.) On the other hand, the parallel data for the Low-Reject condition show that in none of the eleven groups was there a single co-member whose response pattern showed a predominance of support over opposition. In short, these findings suggest that, when an untalkative hint-holder does succeed in getting the elegant solution accepted, he does so because he has the support of a more talkative member—usually the "second man" whom Bales has characterized as frequently a specialist in social-emotional activity in the group.

SUMMARY

An experiment in which 32 four-man groups discussed "human relations problems" provided an opportunity to test a finding of Bales and others that the discussant who talked the most contributed most to the solution of the problem. By providing a hint about a uniquely good, insight solution to one problem, and by differentially locating that hint in the hands of either the highest or the lowest talker in the group, an opportunity was afforded to assess both the perceived and the actual contribution of the top man compared to the bottom man. The results show that top men are almost uniformly perceived as contributing more, and that they are in fact more influential in getting the elegant solution (contained in the hint) accepted by the group. Further data suggest that the differential ability of top men to exert influence is related more to their ability to win attention and support from the group than it is to their ability to reduce opposition. Neither measured intelligence nor fluency and skill in persuasion seem to be important factors in determining the hint-holder's influentiality. Finally, it appears that when highly talkative hint-holders fail to get the elegant solution accepted, it is probably because they are unconvinced of its value and do not advocate it strongly. When untalkative hint-holders succeed in getting acceptance, they do so with the support of one of the more talkative members of the group.

Manuscript received: February 19, 1958

Revised manuscript received: July 11, 1958

Henry W. Riecken

Laboratory for Research in Social Relations

University of Minnesota

Minneapolis 14, Minnesota

REFERENCES

1. Bales, R. F., "The Equilibrium Problem in Small Groups," in T. Parsons, R. F. Bales, and E. A. Shils, *Working Papers in the Theory of Action*, Glencoe, Ill.: The Free Press, 1953.
2. Bass, B. H., "An Analysis of Leaderless Group Discussion," *Journal of Applied Psychology*, 1949, 33, 527-533.
3. Horowitz, M. W., J. Lyons, and H. V. Perlmutter, "Induction of Forces in a Discussion Group," *Human Relations*, 1951, 4, 57-76.
4. Maier, R. F., "The Quality of Group Decisions as Influenced by the Discussion Leader," *Human Relations*, 1950, 3, 155-174.
5. Norfleet, B., "Interpersonal Relations and Group Productivity," *Journal of Social Issues*, 1948, 4, (2), 66-69.

Leadership and Crises

ROBERT L. HAMBLIN, *Washington University*

The purpose of this paper is to report a laboratory investigation¹ of two hypotheses about leadership during crises, namely: leaders have more influence during periods of crisis than during noncrisis periods, and groups tend to replace their old leader with a new leader if the old leader does not have a solution to a crisis problem.

THEORY

Crises

All groups, whether they are large or small, powerful or weak, have the possibility of experiencing a crisis, *an urgent situation in which all group members face a common threat*. A common crisis experienced by family groups is the reduction or loss of income through unemployment, sickness, or death. Religious groups may face crises of persecution. Political parties usually experience a crisis in every election or, if there is lack of electoral machinery, in every revolution. Nations face a crisis in every sudden economic depression or inflation and in every attack by another nation. A crisis is a generic social experience.

There have been a number of field and laboratory studies of group behavior during crises. Stouffer, *et al.*, have studied the reaction of troops to the crisis of battle (20). Durkheim has studied the influence of religious, economic, and political crises on suicide rates and social integration (5). Hovland and Sears (12), Marshall (13), and Moore (15) have studied some social effects of economic crises at the societal level. Also, there have been numerous laboratory studies.² However, only in Marshall's study is there an investigation of the effects of crises on leadership.

Leadership

When most people think of leadership or leaders they think of traits. This may be because of a psychological bias in our culture where aspiring

¹ This research was conducted while the author was on the faculty of Iowa State College. It was supported by funds from Project 262 of the Industrial Science Research Institute of Iowa State College. The Social Science Institute of Washington University furnished editorial and stenographic assistance. An abridged version of this paper was read at the 1958 Annual Meeting of the American Sociological Society at Seattle, Washington. For reports of other aspects of this experiment, see (8) and (10).

² For references to most of these studies, see (8) and (10).

Ben Franklins attempt to develop the traits of a leader as a means of ascending the ladder of success. Consistent with this bias, social scientists have conducted a number of studies in an attempt to isolate the crucial traits of leaders. These attempts have been abortive, not in the sense that each investigator was unable to isolate crucial traits that separated the leaders from the nonleaders, but rather that few commonalities appeared among the crucial traits isolated in the different studies. The brave souls (2, 19) who have tried to synthesize the results of these studies have become discouraged with the trait approach to the study of leadership and have evidently communicated this discouragement quite effectively as most psychologists seem to be abandoning it (4, pp. 535-550; 7; 16, pp. 328-340).

Modern Machiavellis have followed an alternative approach which has been more fruitful. An interest not in what leaders *are* (traits) but in what *successful* leaders *do* has led them to study in experimental situations the effects of a number of different leadership procedures. (The Lewin-Lippitt-White-Autocracy Democracy experiment is, perhaps, the best known of these studies.) Although their substantive findings are important, a significant scientific "breakthrough" was achieved by this group in their development of reliable observational techniques for measuring interaction variables including leadership or influence.

These observational techniques led to the discovery that the leader-follower dichotomy is quite misleading in that influence is almost always distributed, sometimes quite evenly, among group members. That is, the activities of the leader, the high influencer, differ only in degree from the activities of the nonleaders.

This discovery that influence is widely distributed has led in turn to an interest in situational variables which affect the distribution of influence. In particular, the effect of size on the distribution of influence has been investigated several times (1, 9, 11, 14). The present study represents a continuation of this interest in situational variables which affect the distribution of influence. The situational variable here is the "crisis" versus the "noncrisis" situation.

However, there is another important finding from these observational studies which is important in understanding this experiment. After isolating a number of leadership functions, investigators have discovered that a group may have more than one leader at a time. Different members may take the lead in fulfilling different group functions. One member may have the most substantive influence, that is, the most ideas adopted as to how to solve the group's environmental problems. He is called the substantive or task leader. Another member may influence the group the most in coordinating the activities of the various members into a cooperating whole. He is called a procedural

leader. Another group member may have the most influence in helping group members handle their emotions and thus in maintaining group cohesion. He is called the socioemotional leader. (Of course, a single individual may at the same time be the task, procedural, and socioemotional leader.³) In this experiment we are concentrating on the investigation of substantive leadership since the crisis problem is an environmental problem.

The Hypotheses

The first hypothesis (the centralization hypothesis) came from Weber's writings on bureaucracy. Bureaucracies are intricate mechanisms developed to foster the complete centralization of substantive influence in the hands of one man, the director of the bureau. The hundreds and sometimes thousands at the lower levels are supposed to follow the orders of the director in solving the bureau's environmental problems.

Weber assumes that bureaucracies develop when complex problems requiring the coordinated efforts of a vast number of people have to be solved in a very limited time, such as when an army is in battle. Of course, complex problems, coordinating the efforts of vast numbers, and limited time add up to urgency or time pressure. Bureaucracy (and, perhaps, more generally, centralization of influence) occurs because, as Weber points out, "Precision, speed, unambiguity . . . reduction of friction . . . are raised to an optimum point . . ." (6).

As noted above, an essential characteristic of all crises is urgency or time pressure. Hence, from Weber's writings we are led to expect that influence may be more centralized (that the high influencer would have relatively more influence) in crisis than in noncrisis periods.

The second hypothesis (the replacement hypothesis) comes from a working assumption of almost every student of modern politics that a leader (or party) will be voted out of power if he (or it) fails to cope successfully with any serious domestic or international crisis.

Of course, there are some crisis problems for which humans have no known solution. But even in these crises the leader is expected to be replaced. In a study of national politics from 1824 to 1924, Marshall (13) presents rather striking data which support this assumption. From 1824 to 1924 the economy of the United States was predominantly agricultural and agricultural econo-

³ Using data from his interaction categories, together with data from a postexperiment questionnaire, Bales discovered a tendency for one member to become what he called the task leader and another member to become the socioemotional leader. His data indicate that the task leaders were high both on substantive and procedural influence. However, using a set of influence categories which may be more sensitive, we found an increasing tendency for one member to become the substantive leader and another to become the procedural leader as group size increases (9).

mies usually suffer an economic crisis during periods of drought. Marshall's data show that in 11 out of 13 elections where rainfall was below average during the four years prior to election, the party in power was voted out of office. His data also show that in 11 out of 12 elections where rainfall was above average during the four years prior to elections, the political party in power remained in office.

It is a long way from national political parties to the high influencer in a laboratory group and it is a long way from economic crises experienced by nations to an experimentally produced crisis experienced by *ad hoc* groups of three in a laboratory. But if the assumption or principle is general it should apply. Hence, the hypothesis that groups tend to replace their old leader with a new leader if the old leader does not have a solution to the crisis problem.

METHOD

The Experimental Groups

The experiment reported herein involves the before-after observations of twelve three-person groups in a crisis situation and twelve more three-person groups in a control situation. However, this experiment was preceded by two rather elaborate pretests which together involved the observation of an additional 60 groups. It took us this long to work out the measurement problems encountered as well as to acquire the courage to produce a crisis that was severe enough to be valid or, for that matter, interesting.

The participants were either personal acquaintances of the experimental staff or residents of a housing development for married students. Ages ranged between twenty-five and thirty years. Half of the participants were men; the other half women. To reduce barriers to cooperation, participants composing any given experimental group were of the same sex and of approximately the same age.

The Experimental Situation

The experimental task involved a modified shuffleboard game which lasted about 30 minutes.⁴ Each group of participants was ushered into the experimental room, shown the experimental equipment, and given a general but very incomplete idea about the nature and rules of the game. They were told that they were to discover the rules themselves by trying different things and by watching a light board. A red light would flash every time a rule was violated and a green light would flash every time a score was made.

The participants were told that they would be competing with high-school

⁴ For a rather complete set of the instructions used in this experiment see (10).

students who had previously participated in a similar experiment; that their cumulative scores and the average cumulative scores of the high-school students would be posted for each of six, 5-minute playing or task periods into which the game was divided. To ensure their ego involvement the experimenter said:

... the ability to analyze a rather complex situation is the important skill in this game. I have seen many groups work very hard to make many successful shots only to lose most of the points gained through penalties because they failed to learn the ... rules. For this reason we expect mature college graduates to do better than high-school students. If you do get more points, however, you will have to work hard; these high-school students did.

These instructions were evidently successful for the average group behaved as though they were in a tournament. They rushed, ran, and not infrequently shouted. Most groups were visibly satisfied as they mastered the rules and as their scores began to exceed those of the high-school students. The groups were not given the scores they actually earned, however. The threat that would have been experienced by some of the groups who did not do so well as the high-school students had to be avoided. Hence, standard scores were given. These standard scores exceeded the high-school scores by a small but comfortable margin at the end of the first half, or experimental Period I.

By the end of Period I the average group had learned most of the rules and, hence, was making numerous successful shots as indicated by the frequent flashing of the green lights and the before-mentioned lead over the high-school scores. The participants presented a picture of self-satisfaction and confidence. The situation during the last half, or experimental Period II, remained about the same for the control groups. They usually worked hard consolidating their lead over the high-school students and their scores continued to exceed the scores of the high-school students by an ever-widening margin. But the situation changed markedly for the crisis groups. The self-satisfaction and confidence they exhibited at the end of Period I was that "which goeth before a fall."

The Crisis

The crisis was produced by changing the rules of the game. Procedures that were permissible before the change were now against the rules; procedures that were against the rules now became permissible. Lights had been used in teaching the participants the original rules and they could be used in indicating the changes. As the participants saw it, they were receiving red lights for doing the very things for which they had been receiving green lights. But this was not all. As soon as the participants learned a new rule and received a green light, the rule was changed again. The new effect was that

the participants were unable to earn a single score during the last three periods. Since high-school groups were control groups, their scores continued to increase during the last half. Of course, the members of the crisis groups did not know this. As they saw it, their leads vanished, and then their scores fell farther and farther behind those of their rivals.

In the context of the self-satisfaction and confidence at the end of Period I, the switch in Period II to failure of previously successful procedures, the inability to find new procedures that were permanently successful, and the ever-increasing margin of failure in the scores during Period II proved quite effective in producing threat or frustration. The crisis groups looked more frustrated and they engaged in much more aggression than did the control groups during the last half. (The difference in the frequency of observed aggression was significant beyond the .05 level.⁵) Also after the experimental session was over, the crisis groups were very hostile. It took a long, permissive discussion and all the skill the experimenters had to reduce this hostility. However, after discussing the experimental procedures, purpose, and results, most of the participants in the crisis groups acted happy and proud that they had received the "full treatment." Even so, there were some who went away wearing hostile expressions.

But our real fear in this final experiment was not that the threat might not be severe enough but that it would be so severe that the situation would seem hopeless rather than urgent, that the crisis groups would simply withdraw. Not one group gave up, however. Until the end the members of all crisis groups continued to search for, find, and test new procedures. Hence, the situation was evidently both threatening and urgent.

Measurement

As mentioned above, the leadership variable used in this study involves substantive influence or ideas for solving the environmental problems which are adopted by the group.

The environmental problem directly involved in the crisis was to determine the rules by trying different procedures for playing the game and by watching the flashing lights. In attempting to solve this problem, the participants would usually suggest different procedures to be tested. In addition, as the different procedures were tested, the participants would suggest that certain rules were in effect. Not all these suggestions or ideas for solving the environmental problem were adopted by the group. If either or both of the other participants did cooperate in the test of a suggested procedure, a unit of influence was scored. (Most of the suggested procedures were tried and tested by all participants, as the flashing lights were ambiguous enough to make

⁵ For the details of this finding, see (10).

a multiple test desirable.) Also, a unit of influence was scored when a rule was suggested but not immediately tested if there were some agreement, an absence of disagreement, and subsequent conformity to the rule.

An observation form was developed with a column for each participant and a column for "comments." Each time a participant suggested a procedure to the group for testing, or suggested a rule, a "/" was placed in his column in a new row, and a word or two was placed in the same row in the "comments" column to help the observer recall the suggestion if its adoption was delayed. If a suggestion was actually adopted (as indicated above) the "/" was changed to an "X" and was counted as a unit of substantive influence. Scripts were used in training the observers and it was possible to reduce the reliability error to less than 10 per cent.⁶

The basic data were used to calculate two standard measures—influence ratios and acceptance rates.

A given member's influence ratio for a given period represents his raw influence score (the number of accepted substantive suggestions) divided by the average of the raw influence scores of the other group members. For example, if during Period I the high influencer has a raw influence score of 10 adopted suggestions and the other members have scores of 3 and 2 accepted suggestions, then the influence ratio for the high influencer would be 4. He would have four times more influence than the average of the other group members.

A given member's acceptance rate for a given period is simply the proportion of his suggestions that were accepted during that period. For example, if during Period I the high influencer made 15 suggestions and of these 10 were accepted, then his acceptance rate would be .67, indicating that two thirds or 67 per cent of his suggestions were accepted by the group.

A very influential leader probably has both a high influence ratio and a high acceptance rate. But of the two, the influence ratio seems to be the more stable and the more crucial indication of leadership. Hence, the leader or high influencer is always the participant with the highest influence ratio except in the case of a tie. Then the tying participant with the highest acceptance rate is the high influencer.

RESULTS

The Centralization Hypothesis

High influencers have more influence during periods of crisis than during periods of noncrisis.

⁶ A description of the computation of the reliability error as well as the training methods used in reducing the reliability error are found in (10).

As indicated above, influence ratios and acceptance rates are standard measures of influence. Also, the crisis groups experienced a crisis during Period II but the control groups did not. Hence, the most obvious implication of the centralization hypothesis is that during Period II the high influencers in the crisis groups should have higher influence ratios and higher acceptance rates than the high influencers in the control groups.

The mean influence ratios for the high influencers during Period II are 3.3 for the crisis groups and 3.0 for the controls. The difference is in the predicted direction, but the probability is .07 which is greater than .05 where the null hypothesis could be rejected.⁷

The mean acceptance rates for the high influencers during Period II are .44 for the crisis groups and .43 for the controls. The probability of an alpha error exceeds .50.⁸ Hence, the data, when tabulated by experimental periods, do not support the centralization hypothesis.

Of course, the data could also be used to calculate the influence ratios and the acceptance rates for the high influencers during each of the six task periods. Since the task periods 4, 5, and 6 correspond to Period II, another obvious implication of the centralization hypothesis is that during task periods 4, 5, and 6, the high influencers in the crisis groups should have higher influence ratios and higher acceptance rates than the high influencers in the control groups.

The differences in the mean influence ratios shown in Figure 1 support the centralization hypothesis. In addition to their being in the right direction, the probability is sufficiently small to reject the null hypothesis.

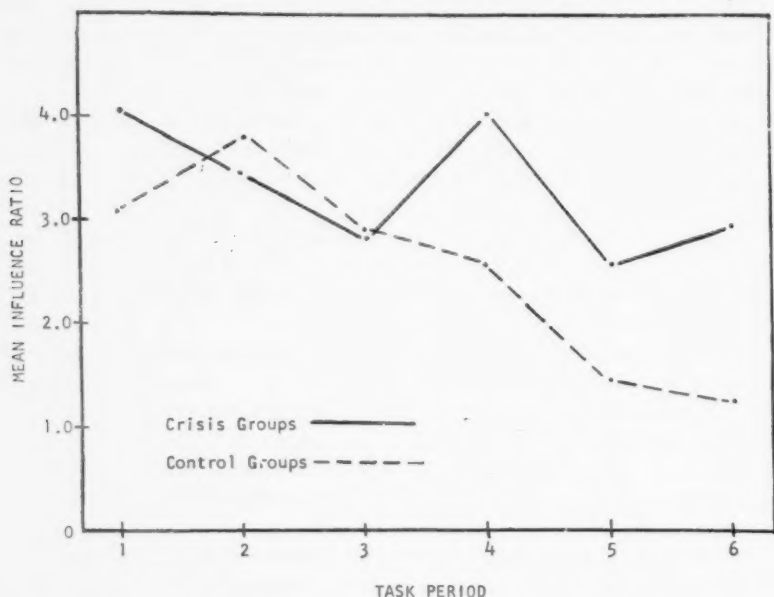
The differences in the mean acceptance rates in Figure 2 support the centralization hypothesis. The differences are in the expected direction for task periods 5 and 6. The probability is sufficiently small to reject the null hypothesis. However, this is not true for period 4. The difference, although it is not significant, is not even in the predicted direction. This indicates that the difference in influence ratios of the high influencers in the crisis and control groups is due to a lack of suggestions on the part of the high influencers in the control groups rather than a decreased suggestibility on the part of their fellow participants.

But why should the data support the centralization hypothesis when tabulated by task period but not support it when tabulated by the longer experimental period? The reason is the replacement of the old leaders by the new leaders which is discussed below.

⁷ This and all the other probabilities of alpha errors (except the one noted in Table 1) were calculated using the Mann-Whitney U test (17, pp. 116-126). The data meet all the assumptions of this distribution-free statistic.

⁸ The term "Type-I error" is sometimes used to refer to an alpha error or the rejection of a null hypothesis when it is in fact true.

FIGURE 1

Mean Influence Ratios for High Influencers in Crisis and Control Groups by Task Periods

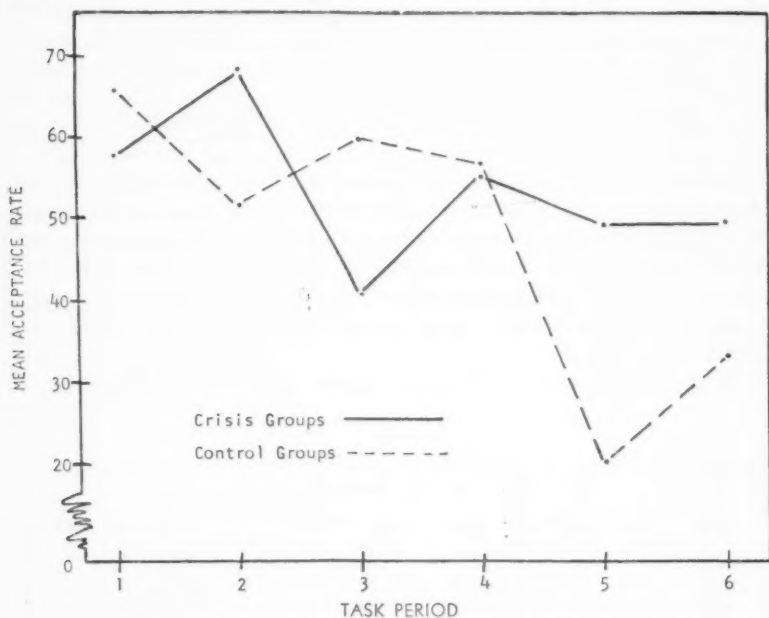
NOTE: Differences between crisis and control groups are significant for periods 4, 5, and 6 (beyond .05 level).

The evidence indicates that replacement does not occur immediately but only after the old leader fails to solve the crisis problem. The old leader usually has a high influence ratio and acceptance rate at first, usually during period 4, but from that point on has a relatively low influence ratio and acceptance rate. The new leader has a low influence ratio and acceptance rate at first but then has a relatively high influence ratio and acceptance rate for the remaining task periods. Thus tabulating the data for the entire experimental period (Period II) gives the old and the new leaders' influence ratios and acceptance rates not for the time each were leaders but for the time they were both leaders and nonleaders. Therefore, tabulating the data by the longer experimental period is inappropriate or, at least, less appropriate than tabulating the data by the shorter task periods.

One last thing before discussing the evidence for the replacement hypothesis. In Figures 1 and 2 notice the high influence ratios and acceptance rates for the first three task periods. They are as high as the influence ratios and

FIGURE 2

Mean Acceptance Rates for High Influencers in Crisis and Control Groups by Task Periods



NOTE: Differences between crisis and control groups are significant for periods 5 and 6 (.05 level).

the acceptance rates of the crisis groups for the last three task periods. Although unintentional, we evidently created crises for both the control and crisis groups for these first three periods. Certainly, in retrospect, a common threat was present (they might not do so well as the high-school students) and the situation was urgent (they had only 30 minutes to learn the rules and make 150 points.) Hence, it may be that the control groups were in a noncrisis situation only during the last three task periods.

The Replacement Hypothesis

A group tends to replace its old leader with a new leader if the old leader does not have an obvious solution to the crisis problem.

Of course the most obvious way to test this hypothesis is simply to count the number of crisis and control groups where the leaders or high influencers did and did not change from Period I to Period II. These data, shown in Table 1, support the replacement hypothesis.

TABLE 1
Replacement of Leaders in the Crisis and Control Groups

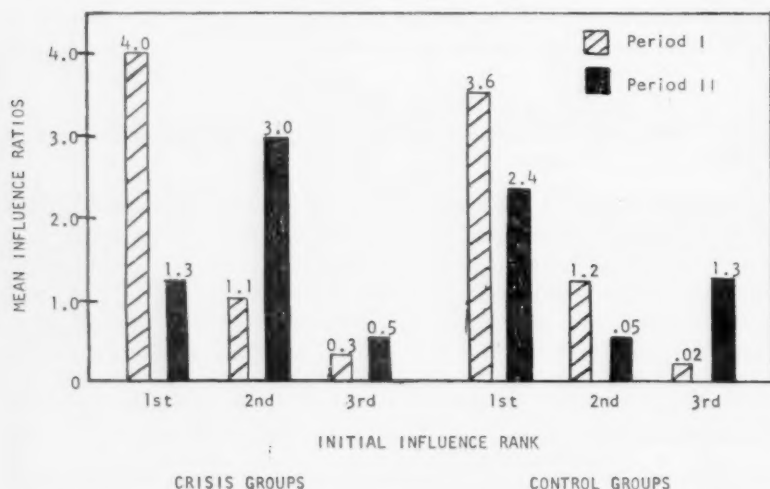
	Leader replaced	Leader not replaced
Crisis groups	9	3
Control groups	3	9

(Chi square, one-tailed test, corrected for continuity, P less than .05)

However, the data in Table 1 tell but a meager part of the story. A more adequate way to show what happened is to tabulate the influence ratios and the acceptance rates of the participants by their initial influence ranks for Period I and Period II. This shows what happened during Period II to the influence ratios and the acceptance rates of the participants who ranked first, second, and third in influence during Period I. These data appear in Figures 3 and 4.

The data in Figure 3 do, of course, support the replacement hypothesis.

FIGURE 3
Mean Influence Ratios by Initial Influence Rank of Members of Crisis and Control Groups by Experimental Periods



NOTE: Change from Period I to Period II is significantly different for participants in crisis and control groups whose initial influence rank was second and third (.05 level).

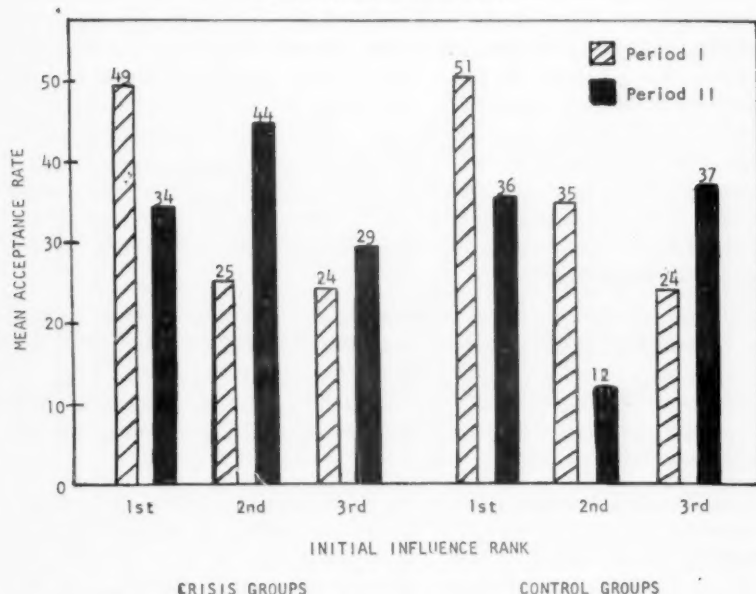
It appears that the participants who initially were second in influence rank became the new leaders in the crisis groups. That is, in fact, what happened.

The participants who were initially second in influence rank became the new leaders in every one of the nine cases where the leader was replaced. This did not happen in the control groups. The participants who were initially second in influence rank either tied for third or became third in influence rank in Period II in ten of the twelve control groups. Also it was the participants who were initially third in influence rank that became the new leaders in the three cases where the leaders were replaced in the control groups.

The mean acceptance rates in Figure 4 tell a similar story except that

FIGURE 4

Mean Acceptance Rates by Initial Influence Rank of Members of Crisis and Control Groups by Experimental Period



NOTE: Change from Period I to Period II is significantly different for participants in crisis and control groups whose initial influence rank was second (.05 level).

the differential increase in acceptance rates of the participants who initially ranked third in influence is not significant.

The data in Figures 3 and 4 combine to suggest that a struggle for influence was waged in both the crisis and control groups. During Period I, the participants who ranked first and second in influence evidently vied with one another for the top position. In the crisis groups the crisis evidently caused

the participants who originally ranked second to become dominant. The participants who ranked third remained quite isolated. In the control groups, however, the participants who initially ranked first were able to maintain their position of dominance evidently by forming a coalition with the participants who initially ranked third, perhaps, by giving them support. This evidently froze out the participants who initially ranked second.

These results are somewhat at variance with those of Mills (14). Mills concluded that three-person groups normally develop into a coalition of two and an isolate, and that Simmel's (18, pp. 165-178) *tertius gaudens* (an influence structure where two powerful units are vying with one another for the support of a weak third unit) is rare. Unfortunately, we do not have the support and opposition data to determine exactly what the coalition structure was. But it is rather obvious that the control groups do exemplify the *tertius gaudens* and, in three cases, the *tertius*, initially the weakest of the three units involved, became the high influencer. It is also very doubtful that the participants initially ranked first and second in the crisis groups actually formed a coalition, at least in the nine cases where replacement occurred. However, in the three cases where replacement did not occur, a coalition may have existed between the participants who initially ranked first and second.

Of course, our experimental situation differs considerably from Mills'. There were a number of factors present which could, conceivably, account for the differential results.

1. *Time pressure.* Our situation was designed to be urgent, and, in general, it was probably much more urgent than was Mills' situation.
2. *Threat.* It was almost entirely absent from Mills' experiment.
3. *Objectivity.* In our experiment, the flashing lights gave the participants rather objective criteria for making decisions. This, of course, was not true in Mills' experiment.

At this point it is impossible to say which, if any, of these factors were responsible for the differential results. Further experimentation is required. However, these results do highlight the importance of attempts to isolate the conditions which produce different influence and coalition structures. Theory such as that developed by Caplow (3) and experiments such as the one conducted by Vinacke and Arkoff (21) are indicated by the results.

SUMMARY

The purpose herein was to report a laboratory investigation of two hypotheses about leadership during crises, namely: leaders have more influence during periods of crisis than during non-crisis periods, and groups tend to replace their old leader with a new leader if the old leader does not have an obvious solution to a crisis problem. Twenty-four groups were brought into a labora-

tory situation. Twelve experienced an apparently genuine crisis where there was no solution to the crisis problem. Data from an observational measure of influence give support to both hypotheses.

Manuscript received: February 3, 1958

Revised manuscript received: June 13, 1958

Robert L. Hamblin
Social Science Institute
Washington University
St. Louis 5, Missouri

REFERENCES

1. Bales, R. F., "The Equilibrium Problem in Small Groups," in T. Parsons, R. F. Bales, and E. A. Shils (eds.), *Working Papers in the Theory of Action*, Glencoe, Ill.: Free Press, 1953, 111-162.
2. Bird, C., *Social Psychology*, New York: Appleton-Century, 1940.
3. Caplow, T., "A Theory of Coalitions in the Triad," *American Sociological Review*, 1956, 21, 489-493.
4. Cartwright, D., and A. Zander, *Group Dynamics*, Evanston, Ill.: Row, Peterson, 1956.
5. Durkheim, E., *Suicide*, Glencoe, Ill.: Free Press, 1951.
6. Gerth, H. H., and C. W. Mills, *From Max Weber: Essays in Sociology*, New York: Oxford University Press, 1946.
7. Gouldner, A. W. (ed.), *Studies in Leadership*, New York: Harper, 1950.
8. Hamblin, R. L., and J. A. Wiggins, "Suggestibility, Imitation, and Recall During a Crisis," *Midwest Sociologist*, 1957, 20, 26-32.
9. Hamblin, R. L., "An Experimental Study of the Relationship of Communication, Power Relations, Specialization, and Social Atmosphere to Group Size," Ph.D. Dissertation, University of Michigan (Mic. 55-1362).
10. Hamblin, R. L., "Group Integration During a Crisis," *Human Relations*, 1958, 11, 67-76.
11. Hare, P., "A Study of Interaction and Consensus in Different Sized Groups," *American Sociological Review*, 1952, 17, 261-267.
12. Hovland, C. I., and R. R. Sears, "Minor Studies in Aggression: VI, Correlations of Lynchings with Economic Indices," *Journal of Psychology*, 1949, 9, 301-310.
13. Marshall, R., "Precipitation and Presidents," *The Nation*, 1927, 124, 315-316.
14. Mills, T. M., "Power Relations in Three-Person Groups," *American Sociological Review*, 1953, 18, 351-356.
15. Moore, B., "A Comparative Analysis of Class Struggle," *American Sociological Review*, 1945, 10, 21-37.
16. Sanford, F. H., "The Followers Role in Leadership Phenomena," in G. E. Swanson, et al., (eds.), *Readings in Social Psychology*, New York: Holt, 1952.
17. Siegel, S., *Nonparametric Statistics for the Behavioral Sciences*, New York: McGraw-Hill, 1956.
18. Simmel, G., *The Sociology of Georg Simmel*, K. H. Wolff (ed.) Glencoe, Ill.: Free Press, 1950.
19. Stogdill, R. M., "Personal Factors Associated with Leadership: A Survey of the Literature," *Journal of Psychology*, 1948, 25, 35-71.
20. Stouffer, S. A., et al., *The American Soldier: Combat and Its Aftermath*, Princeton, N. J.: Princeton University Press, 1949.
21. Vinacke, W. E., and A. Arkoff, "An Experimental Study of Coalitions in the Triad," *American Sociological Review*, 1957, 22, 406-414.

